Exploratory analysis of internet search trends during the COVID-19 outbreak

Análisis exploratorio de las tendencias de búsqueda en internet durante el brote de la COVID-19

Christian Renzo Aquino-Canchari1* https://orcid.org/0000-0002-7718-5598
Sarai Gloria Chávez-Bustamante2 https://orcid.org/0000-0002-8268-9424
Brenda Sofía Caira-Chuquineyra3 https://orcid.org/0000-0003-4787-5552

1Universidad Peruana Los Andes, Facultad de Medicina Humana. Sociedad Científica de Estudiantes de Medicina Los Andes (SOCIEMLA). Huancayo, Perú.
2Universidad Continental. Facultad de Medicina Humana. Sociedad Científica de Estudiantes de Medicina Continental (SOCIMEC), Huancayo, Perú.

*Correspondence: christian.aquino.canchari@gmail.com

ABSTRACT

Coronavirus disease 2019 has put the world in a health emergency. Searching for information on the Internet largely reflects people's interest in this pandemic.

Objective: Conduct an exploratory analysis of Internet search trends during the 2019 coronavirus disease outbreak.

Methods: Google Trends was used to provide data on the relative volume of Google searches for terms related to 2019 coronavirus disease. The evaluation period was from January 01 to May 17, 2020.
Results: The search term used to know this pandemic was “coronavirus”, the most searched symptom was “fever”, followed by “sore throat” and “cough”, in addition, the interest of users to know the transmission routes of the acute respiratory syndrome coronavirus 2. As for preventive measures, the most searched term was “stay home”, followed by “facial masks”, “social distancing” and “washing hands”.

Conclusions: The results confirmed interest in COVID-19 via Internet. Using information from people's Internet search interest could help formulate health policies to better control the 2019 coronavirus disease outbreak.

Key words: COVID-19; coronavirus; SARS-CoV-2; Google Trends; risk communication; pandemic; public health.

RESUMEN
La enfermedad del coronavirus 2019 ha puesto al mundo en una emergencia sanitaria. La búsqueda de información en Internet refleja en gran medida el interés de la gente por esta pandemia.

Objetivo: Realizar un análisis exploratorio de las tendencias de búsqueda en Internet durante el brote de la enfermedad por coronavirus de 2019.


Resultados: El término de búsqueda utilizado para conocer esta pandemia fue “coronavirus”, el síntoma más buscado fue “fiebre”, seguido de “dolor de garganta” y “tos”; además, el interés de los usuarios por conocer las vías de transmisión del síndrome respiratorio agudo del coronavirus 2. En cuanto a las medidas preventivas, el término más buscado fue “quedarse en casa”, seguido de “máscaras faciales”, “distanciamiento social” y “lavarse las manos”.

Conclusiones: Los resultados confirmaron el interés en el COVID-19 a través de Internet. El uso de información del interés de búsqueda de las personas en Internet podría ayudar a formular políticas de salud para controlar mejor el brote de la enfermedad del coronavirus del año 2019.
Palabras clave: COVID-19; coronavirus; SARS-CoV-2; tendencias de Google; comunicación de riesgos; pandemia; salud pública.

Recibido: 18/06/2020
Aceptado: 11/09/2020

Introduction

On December 31, 2019, a total of 41 cases of pneumonia of unknown etiology were confirmed in Wuhan City, Hubei Province, China. Wuhan City is an important transportation hub for China, with a population of over 11 million people. These patients had in common a history of having visited the wholesale seafood and wildlife market, in which animals such as poultry, bats, groundhogs and snakes are also traded.(1)

The causal agent of this new outbreak was quickly identified by Chinese authorities on January 7, 2020, as a new type of virus in the Coronaviridae family, temporarily named as the Novel coronavirus (2019-nCoV). On Thursday, January 30, 2020, the World Health Organization (WHO) declared the infection as a Public Health Emergency of International Concern (PHEIC).(2)

Subsequently, WHO will officially name the disease as "COVID-19"(3) and the International Committee on Virus Taxonomy will name the virus as "severe acute respiratory syndrome coronavirus 2" (SARS-CoV-2).(4) To date, May 2, 2020, there are reported to be 6 318 040 confirmed cases and 376 885 deaths from the spread of COVID-19 in at least 215 countries.(5)

In the last decades, Internet has become more available and popular around the world and constitutes a source of quick access to all kinds of information, with the search for information on Health and Medicine being one of the most consulted content. The Internet and its search engines, including Google, provide us a variety of health-related medical...
websites and portals, and on certain occasions, induce a substitute for diagnosis and treatment by the healthcare professional.\(^6,7\)

Google Trends\(^\text{TM}\) (GT\(^\text{TM}\)) is the most popular tool to examine the changes in interest of the world population about different fields, being health and medicine one of the most consulted. GT\(^\text{TM}\) studies on health research are classified into four areas: infectious diseases, mental health, non-communicable diseases and general population behavior.\(^8\)

This platform allows you to evaluate and compare search terms in Google, according to geographic and temporal patterns, and presentation through specific search volumes (RSV).\(^9\)

Ciaffi et al.\(^{10}\) and Sousa-Pinto et al.\(^{11}\) reported a linear increase between search trends for terms related to SARS-CoV-2 infection in GT with the interest of the media and opinion public on the pandemic in European countries.

The emergence of COVID-19 has gained wide coverage in public media and global news; therefore, search terms related to COVID-19 reflect a worldwide interest in the current pandemic, in addition to concern and fear of infection that exists among the population, then this type of information should be investigated. The objective of this research was to determine Internet search trends during the COVID-19 outbreak.

**Methods**

A retrospective, descriptive and observational study was carried out. GT\(^\text{TM}\) is an online tracking system that allows you to determine the proportion of searches for terms consulted by the user among all searches performed with Google. GT\(^\text{TM}\) normalizes search data to facilitate comparisons between terms, providing a relative search volume (RSV) adjusted for a given location and time period; the values obtained are scaled in a range from 0 to 100 depending on the proportion. Each point on the graph generated by GT\(^\text{TM}\) is divided by the highest point, which is conventionally set to 100.\(^{12}\)

An exploratory analysis of Google's relative search volume was performed via GT (https://trends.google.com/trends/), spanning a “time” period from January 01 to January
17, 2020 choosing “all categories” and “web search”. The search was performed on May 19, 2020. Also, the search and comparison terms were divided into 4 areas: definition of COVID-19; infection mode; symptoms and preventive measures:

1. **Definition of COVID-19**: coronavirus; covid-19; covid; SARS-CoV-2; novel coronavirus.
2. **COVID-19 symptoms**: covid-19 symptoms; cough; fever; shortness of breath; fatigue; sore throat; muscle pain; dysgeusia; anosmia.
3. **COVID-19 infection mode**: coronavirus infection; coronavirus spread; covid spread; droplets; surface transmission.
4. Precautionary measures: wash hands, social distancing, face masks; cover your cough, stay home.

The study did not require the approval of an ethics committee since the data is publicly available, anonymous, and cannot be traced back to identifiable individuals. The results provided by GT were downloaded in standardized values separated by commas (CSV), then extrapolated to a Microsoft Excel spreadsheet in its 2019 version. The quality control of the information was carried out by double data entry, correcting the inconsistencies consulting the originals.

**Results**

About the search terms “coronavirus”; “covid 19”; “covid”; “SARS CoV 2” and “novel coronavirus”, the term with the highest RSV found was “coronavirus”, followed by “covid”, “covid 19”, “novel coronavirus” and “SARS-CoV-2” (Fig. 1).
The search term “coronavirus” was the most searched for in Kiribati, Nauru, Germany, and the Netherlands, while the term “covid-19” was the most searched for in Serbia, Vietnam, Indonesia, and Timor-Leste, in addition to the term “covid” was the most sought for in Iceland, Thailand and El Salvador, in relation to the term “SARS-CoV-2” was the most searched for in Malawi, Eritrea and Liberia. Finally, the term “novel coronavirus” was the most searched for in China, Hong Kong, and Macao (Fig. 2).
**Fig. 2** - Internet search trends by region for the terms “coronavirus”, “covid-19”, “covid”, “SARS-COV-2” and “new coronavirus”.

In the comparative RSV of search terms on COVID-19 symptoms, the term with the highest RSV was “fever”, followed by “sore throat”, “cough”, “covid 19 symptoms” (Fig. 3).

**Fig. 3** - Comparative relative search volume in relation to Internet search trends on COVID-19 symptoms.
The search term “cough” was the most searched in the United Kingdom, India and Ireland, while the term “fever” was the most searched in Mexico, Germany and Indonesia, compared to the term “shortness of breath” was the most searched in the United States and South Africa. Besides, the term “fatigue” was the most searched in France and Sweden, in relation to the term “sore throat” was the most sought after in Nigeria and the UK; then the term “muscle pain” was the most searched in the Netherlands and Saudi Arabia. Furthermore the term “dysgeusia” was the most searched in Indonesia and Japan, and finally the term “anosmia” was the most searched in Spain and Mexico (Fig. 4).

Fig. 4 - Internet search trends by region for the terms “covid-19 symptoms”, “cough”, “fever”, “shortness of breath” and “fatigue”.

Regarding the comparative RSV of search terms in the SARS-CoV-2 transmission routes, the term with the highest RSV was "coronavirus spread", followed by “coronavirus infection”, “covid spread”, “droplets” and “surface transmission” (Fig. 5).
The search term “coronavirus infection” was the most searched in France and South Korea, while the term “coronavirus spread” was the most searched in Italy and Sweden, in addition to the term “covid spread” was the most searched in Malaysia and the Philippines. Then, the term “droplets” was the most searched in Indonesia and the Philippines. Finally, the term “surface transmission” was most searched in France and Ireland (Fig. 6).
Concerning the comparative RSV of search terms on preventive measures of COVID-19, the term with the highest RSV was “stay home”, followed by “face masks”, “social distancing”, “wash hands” and “cover your cough” (Fig. 7).

**Fig. 6** - Internet search trends by region for the terms “coronavirus infection”, “coronavirus spread”, “covid spread”, “droplets” and “surface transmission”.

**Fig. 7** - Comparative relative search volume in relation to Internet search trends on measures to prevent COVID-19.
The search term “wash hands” was the mostsearched in Spain and New Zealand, while the term “social distancing” was the most searched in the Philippines and Australia, compared to the term “face masks” was the most searched in South Africa and Ireland. Then, the term “cover your cough” was the most searched for in New Zealand and The United States, and finally, the term “stay home” was the most searched for in Malaysia and India (Fig. 8).

**Discussion**

The present study evaluates the Internet search attentions on COVID-19. Due to the pandemic, millions of people remain in their homes as a preventive measure to reduce contagion.\(^{(12)}\) Web-based research has increased in the past decade. Several studies exposed the utility of the GT\(^{\text{TM}}\) to know the need for information and estimate the incidences in different diseases (for example, osteoarthritis, breast cancer or COPD).\(^{(13,14,15)}\) In addition, infodemiology and GT\(^{\text{TM}}\) are used to generate awareness profiles and can be a substitute for data collection, such as surveys.\(^{(16)}\)

Our study showed that the term with the highest RSV was “coronavirus”, which peaked on March 15 worldwide, a date that coincides with a total of 142 000 confirmed and, in addition, the first cases were imported into Latin America.\(^{(17)}\) Kiribati, the country where
the term “coronavirus” was the most searched, to date does not present any case of coronavirus according to the panel of “Johns Hopkins University Center for Systems Science and Engineering Dashboard”. This could be due to the fact that the population has been kept informed and prepared.

The worldwide search for the term “SARS-CoV-2” was not considered very important until January 21, when the term was first associated with the development of the virus, reported by the WHO in the United States, and then had a constant daily minimum search, this could be due to the complexity of the term. Furthermore, this study conducted from December 31 to February 24, 2020 showed that on January 31 the term “COVID-19 and its variants” had the highest search peak in the world. Another study reveals that on March 12 it had a maximum public interest in the subject and was due to the fact that on March 11, the WHO classified the coronavirus outbreak as a pandemic; date that coincides with our third peak of maximum search.

Respiratory symptoms are extremely variable, from minimal symptoms to hypoxia significantly with ARDS. According to K. Yuki et al., The time between the onset of symptoms and the development of ARDS can be up to 9 days, which can mean that respiratory symptoms can progress rapidly. According to the WHO, the most common symptoms of COVID-19 include: “fever”, “dry cough” and “fatigue”. On the other hand, there are less frequent symptoms, such as: “sore throat”, “muscle pain”, “dysgeusia”, “anosmia”, “shortness of breath”, among others. The term with the highest amount of RSV in terms of symptoms was “fever” and this could be due to the fact that it is the most frequent symptom in infected patients.

In our study, the term “dysgeusia” was the least searched for among the symptoms and may be due to its infrequency. In Wuhan was found that the search term dysgeusia increased when new cases were reported, serving as a triage symptom for the disease, on the other hand, the term “anosmia” was also poorly received, but this is considered relevant symptom in at least 8 countries. Although it is true, the search for symptoms is related to other pathologies; however, a larger increase in terms has been observed during this period of the pandemic.
The National Health Commission of the People's Republic of China recognizes that the main routes of transmission of COVID-19 are mainly through by direct contact through respiratory droplets, which occurs when the infected person coughs or sneezes, and indirect contact, it happens when a person touches a surface or object contaminated with the virus and then touches his mouth, nose, or eyes. Another form of SARS-CoV-2 infection can occur in specific places and circumstances where procedures are performed or medical treatments that can produce aerosols are administered (for example, endotracheal intubation, bronchoscopy, open aspiration, dental care, among others).(25) There has been some evidence that SARS-CoV-2 can cause intestinal infection and be present in feces, however, only one study has cultured this virus from a single stool sample, and to date, no fecal-oral transmission has been reported.(26)

Our study shows that the search term “coronavirus spread” had the highest RSV compared to the others. In addition, it showed an increase in searches on January 21, 2020, the date that Washington state officials confirmed the first case of COVID-19 in the US, being the first in the American continent.(27)

Search trends in contact routes inform us of an increase in the search for “coronavirus spread” during the week of March 07, which could explain some opinions and informed decisions during these dates, the peak of searches was the March 12, this may explain that the previous day, the WHO director-general expressed his deep concern about the alarming levels of spread and severity, transferring that in two weeks prior to that week, the number of infected in all the world outside of China had multiplied by 13 and the number of affected countries had tripled.(25) Therefore, characterized COVID-19 as a pandemic.

Prevention strategies against COVID-19 are public health measures that allow delaying or preventing the transmission of the infection, these were decisive in controlling the SARS epidemic in 2003. In the absence of a vaccine against COVID-19, the best way to prevent is avoiding exposure to the virus through various measures, such as the use of face masks; cover coughs and sneezes with disposable tissues or, in any case, use your flexed elbow;
wash your hands regularly with soap and water, or use a hand sanitizer that contains at least 60% alcohol; avoid contact with infected people; maintain an adequate distance (greater than 1 meter) and avoid touching your eyes, nose and mouth with your hands without washing.\(^{(28)}\)

Regarding prevention measures, our study showed that the search term "stay home" had the highest RSV. The stay-at-home order is part of quarantine, one of the oldest and most effective tools to control the spread of infectious outbreaks. As reported by Lyu et al., it is suggested that the order to stay home in the state of Iowa helped decreasing the number of cases compared to the border state of Illinois, where this measure was not implemented;\(^{(29)}\) which is consistent with the findings of Friedson et al., where the order to stay home issued on March 19, was associated with a reduction in cases by COVID-19 one month after the order.\(^{(30)}\)

We observed an increasement in the search to «stay home» during the week of March 23, which could explain some measures implemented by the WHO, such as WhatsApp health alert launched on March 20 and the joint awareness campaign with the Federation of International Football Associations (FIFA), launched on March 23, where world-renowned footballers demand that the world population take the five key steps to stop the spread of COVID-19.\(^{(31)}\)

The search term “face masks” showed a relative increase in search trends. When COVID-19 became a global health emergency, there was a visible contrast between the responses of citizens in East Asia and the rest of the world. In countries such as China, Japan, South Korea, among others, the use of masks was general and sometimes mandatory.\(^{(32)}\) In Europe and North America, concerned citizens were repeatedly told that masks are not recommended for general use. However, with the increase in cases, now, an increasing number of agencies and governments, including the Czech Republic and the US CDC. And the latter urges general population to wear masks, but others, such as the WHO and Public Health of England do not do so because they recommend the use of masks by medical personnel and patients who are ill or who care for a COVID-19 patient at home.\(^{(33)}\)
Our findings indicate that GT can be used as a tool to monitor people's concern for COVID-19, predict future outbreaks and epidemics, complementing traditional public health surveillance systems.

**Conclusions**

The results obtained confirm the interest in COVID-19 through the internet. The search term used to know this pandemic was “coronavirus”, the most searched symptom was “fever”, followed by “sore throat” and “cough”, in addition, the interest of the users to know the transmission routes was evidenced of the SARS-CoV-2. As for preventive measures, the most searched term was “stay home”, followed by “facial masks”, “social distancing” and “washing hands”.

In recent decades, the use of the Internet has increased. Google Trends emerges as a complementary tool for the epidemiological surveillance of diseases, especially in countries with limited resources, being an option for making timely decisions, especially in the case of diseases prone to epidemics. Future studies should focus on predicting outbreaks and epidemics by using mathematical models that fit other parameters.

**Limits**

1. In view of the fact that GT only provides data based on the measure of "interest", more accurate and informative models could be constructed if the absolute search frequency were available to researchers.
2. The keywords selected in the study may not identify the interest of Google users in relation to COVID-19.
3. It is recommended that future researchers combine other sources of data, such as information from social networks, contacts of people with the special call center for COVID-19, the media, among others.
Bibliographic references


Conflicts of interest
There authors declare no conflicts of interest.

Authors contributions
Christian Renzo Aquino-Canchari: Participated in conception and design of study, data collection, analysis and interpretation of data, critical revision, drifting and reviewed to the final manuscript.
Sarai Gloria Chávez-Bustamante: Participated in desing of study, data collection, analysis and interpretation of data, drifting and reviewed to the final manuscript.
Brenda Sofía Caira-Chuquineyra: Participated in desing of study, data collection, drifting and reviewed to the final manuscript.

Funding
Self funding.