An Empirical Infodemiology Study of What People Learned about COVID-19 and Behavior Towards Public Health Guidelines Using Web Searches

Ikpe Justice Akpan, Obianuju Genevieve Aguolu, Yawo Mamoua Kobara, Rouzbeh Razavi, Asuama A. Akpan, Murali Shanker

Submitted to: Journal of Medical Internet Research
on: May 01, 2021

Table of Contents

An Empirical Infodemiology Study of What People Learned about COVID-19 and Behavior Towards Public Health Guidelines Using Web Searches

Ikpe Justice Akpan[1*] BSc, MA; Obianuju Genevieve Aguolu[2*] MD, MPH, PhD; Yawo Mamoua Kobara[3*] BSc, MSc; Rouzbeh Razavi[4*] BSc, MSc, PhD; Asuama A. Akpan[5*] BSc, MA; Murali Shanker[4*] BSc, MSc, PhD

[1]Department of Management & Information Systems, Kent State University New Philadelphia US
[2]Infectious Disease Internal Medicine Department, Yale School of Medicine Yale University New Haven US
[3]Statistical and Actuarial Sciences Western University London CA
[4]Department of Management & Information Systems Kent State University Kent US
[5]Research and Development, Ibom International Center for Research and Scholarship WINDSOR CA
[*]these authors contributed equally

Corresponding Author:
Ikpe Justice Akpan BSc, MA
Department of Management & Information Systems, Kent State University
330 University Drive Northeast
New Philadelphia
US

Abstract

Background: The use of the Internet and web-based platforms to obtain public health information and manage health-related issues has become widespread in this digital age. The practice is so pervasive that the first reaction to obtaining health information is to 'google it.' As SARS-CoV-2 broke out in Wuhan, China, in December 2019, and quickly spread worldwide, people flocked to the Internet to learn about the novel coronavirus and the disease, COVID-19. Lagged response by governments and public health agencies to prioritize the Internet and the World Wide Web to disseminate information about the coronavirus outbreak and building trust gave room for others to quickly populate the social media, online blogs, news outlets, and websites with misinformation and conspiracy theories about the COVID-19 pandemic resulting in people's deviant behaviors towards public health safety measures.

Objective: This study investigates what people learned about COVID-19 through "web search," exposure to misinformation and conspiracy theories, and the impacts on behaviors towards public health safety measures.

Methods: We used the Google trends worldwide search index covering the first six months after the SARS-CoV-2 outbreak (January to June 2020) when the public scrambled for information about the pandemic. Data analysis employed statistical trends, correlation and regression, principal component analysis, and predictive models.

Results: (i). The principal components analysis identifies two latent variables comprising past coronavirus epidemics (pastCoVepidemics) and the ongoing COVID-19 pandemic (presCoVpandemic). Both principal components (PCs) were utilized significantly to learn about SARS-CoV-2 and COVID-19 and explained 88.78% variability. (ii). Three (3) PCs fuelled misinformation about COVID-19 [Misinformation("Biological Weapon," "VirusHoax," "common cold," "COVID-19Hoax," "ChinaVirus"); ConspTheory1("@5G"); ConspTheory1("IngestBleach")]. These PCs explained 84.85% of the variability, (iii). Two (2) PCs identified two components of public health measures [PubHealthMes1("Social Distancing," "WashHand," "Isolation," "Quarantine."); PubHealthMes2("WearMask"), which explained 84.7% of the variability. (iv). Based on the PCA results, log-linear, and predictive models [ConspTheory1(@5G)] are identified as a predictor of people's behavior towards public health measures (PubHealthMes2). Although (r=0.83), (r=-0.11) for Misinformation(COVID-19Hoax, VirusHoax, common cold, and more) and ConspTheory2(Ingestbleach), respectively, with PubHealthMes1(social dist, handwash, isolation, and more), both were not statistically significant with (p=0.267), (p=0.13), respectively.

Conclusions: Several studies focus on the impacts of social media and related platforms on spreading misinformation and conspiracy theories. This study provides the first empirical evidence to the mainly anecdotal discourse on the use of web search to learn about SARS-CoV-2 and COVID-19.

(JMIR Preprints 01/05/2021:28975)
DOI: https://doi.org/10.2196/preprints.28975

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

An Empirical Infodemiology Study of What People Learned about COVID-19 and Behavior Towards Public Health Guidelines Using Web Searches

Ikpe Justice Akpan[1*], PhD; Obianuju Genevieve Aguolu[2], MBBS, MPH, PhD; Rouzbeh Razavi[1], PhD; Yawo Mamoua Kobara[3], BS, MSc; Asuama A. Akpan[4], BSc Ed., MED; Murali Shanker[1], PhD

1. Department of Management & Information Systems, Kent State University, Ohio, USA.

2. Infectious Disease Internal Medicine Department, Yale School of Medicine, Yale University, New Haven, Connecticut, USA.

3. Department of Statistics and Actuarial Sciences, University of Western Ontario, London ON, Canada.

4. Health Informatics Research Group, Ibom International Center for Research and Scholarship, Windsor ON, Canada.

***Corresponding Author**

Ikpe Justice Akpan, Ph.D.
Department of Management & Information Systems
Kent State University,
330 University Drive NE,
New Philadelphia, OH 44663
USA
Phone: +1 330-339-3391 EXT 7572
Email: iakpan@kent.edu

An Empirical Infodemiology Study of What People Learned about COVID-19 and Behavior Towards Public Health Guidelines Using Web Searches

Abstract

Background: The use of the Internet and web-based platforms to obtain public health information and manage health-related issues has become widespread in this digital age. The practice is so pervasive that the first reaction to obtaining health information is to 'google it.' As SARS-CoV-2 broke out in Wuhan, China, in December 2019, and quickly spread worldwide, people flocked to the Internet to learn about the novel coronavirus and the disease, COVID-19. Lagged response by governments and public health agencies to prioritize the Internet and the World Wide Web to disseminate information about the coronavirus outbreak and building trust gave room for others to quickly populate the social media, online blogs, news outlets, and websites with misinformation and conspiracy theories about the COVID-19 pandemic resulting in people's deviant behaviors towards public health safety measures.

Objectives: To determine what people learned about the COVID-19 pandemic through "web search;" examine any association between what people learned about COVID-19 and behavior towards the public health guidelines; and analyze the impact of misinformation and conspiracy theories about the COVID-19 pandemic on people's behavior towards public health measures.

Methods: This paper undertakes an infodemiology study using the Google trends worldwide search index covering the first six months after the SARS-CoV-2 outbreak (January 1st, to June 30th, 2020) when the public scrambled for information about the pandemic. Data analysis employed statistical trends, correlation and regression, principal component analysis, and predictive models.

Results: (i). The principal components analysis identifies two latent variables comprising past coronavirus epidemics (pastCoVepidemics) and the ongoing COVID-19 pandemic (presCoVpandemic). Both principal components (PCs) were utilized significantly to learn about SARS-CoV-2 and COVID-19 and explained 88.78% variability. (ii). Three (3) PCs fuelled misinformation about COVID-19 [Misinformation("Biological Weapon," "VirusHoax," "CommonCold," "COVID-19Hoax," "ChinaVirus"); ConspTheory1("@5G"); ConspTheory1("IngestBleach")]. These PCs explained 84.85% of the variability, (iii). Two (2) PCs identified two components of public health measures [PubHealthMes1("Social Distancing," "WashHand," "Isolation," “Quarantine.”); PubHealthMes2(“WearMask”), which explained 84.7% of the variability. (iv). Based on the PCA results, log-linear, and predictive models [ConspTheory1(@5G)] is identified as a predictor of people’s behavior towards public health measures (PubHealthMes2). Although (r=0.83), (r=-0.11) for Misinformation(COVID-19Hoax, VirusHoax, CommonCold, and more) and ConspTheory2 (Ingestbleach), respectively, with PubHealthMes1(social dist, handwash, isolation, and more), both were not statistically significant with (p=0.267), (p=0.13), respectively.

Conclusions: Several studies focus on the impacts of social media and related platforms on spreading misinformation and conspiracy theories. This study provides the first empirical evidence to the mainly anecdotal discourse on the use of web search to learn about SARS-CoV-2 and COVID-19.

Keywords: Internet, novel coronavirus, SARS-CoV-2, COVID-19, infodemiology, misinformation, conspiracy theories, public health

Introduction

A novel (n) coronavirus (CoV) initially named 2019-nCoV emerged in Wuhan, China, and was formerly reported to the World Health Organization (WHO) on December 31, 2019 [1-3]. Further scientific evidence soon unveiled the semblance of the 2019-nCoV's genome sequence and a previous epidemic, the "severe acute respiratory syndrome" (SARS), a disease epidemic caused by SARS coronavirus (SARS-CoV), which broke out in Foshan, China, in 2002 [4,5]. Some initial studies also identified similar features that related to the

"Middle East respiratory syndrome" (MERS) epidemic caused by the MERS coronavirus (MERS-CoV) as the causative agent [6].

The outbreak was formerly named 2019 novel coronavirus (2019-nCoV) on January 13, 2020, the same day that the first imported case outside China occurred in the Philippines and other countries [7]. The spread of 2019-nCoV continued across many countries, causing the WHO to declare the outbreak a pandemic [6]. The 2019-nCoV was later renamed SARS-CoV-2 and identified as the causative agent of the coronavirus disease 2019 (COVID-19) in February 2020 [8,9]. The highly contagious COVID-19 spread rapidly globally and caught the world unprepared. With no adequately planned health communication strategies, panic ensued, while confirmed cases of infections and deaths from COVID-19 increased rapidly worldwide [3]. The public rushed to Internet platforms to learn about the outbreak through Google search, online news outlets, and social media platforms [10-15].

In March 2020, the WHO launched a free online introductory training course in different languages (including English, French, Spanish, and Chinese) to make the public aware of the contagious COVID-19 [6]. However, it is unclear how many people knew or utilized the free training lessons about COVID-19 that WHO had made available via its website [6]. Instead, several studies suggest that the public mflocked to the Internet to learn about SARS-CoV-2 and COVID-19 through web searches, Online news outlets, and social media [16,17]. Analyzing how people search and navigate the World Wide Web and other Internet platforms for health-related information can provide valuable insights into the health-related behavior of populations [18-20]. The public's preference for online health information closely matches the field of Infodemiology, a term that is a portmanteau of information and epidemiology. According to Eysenbach [18], the term is defined as the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the aim to inform public health and public policy. Considering the global spread of COVID-19, using the Internet to learn or gain information about the pandemic in this digital age is not surprising as Internet use has become pervasive worldwide [21,33]. Several studies have examined social media's influence on what people learned and the appropriate behaviors towards misinformation and conspiracy theories [14,15]. Similarly, Sulyok et al. [12] examined the impact of web search on the confirmed cases of COVID-19 in Europe, while Neely et al. [22] investigated information-seeking behaviors on social media for the pandemic.

Miller et al. [14,15] identify political leaders' failure in sensitizing the public as a motivating factor that pushes people to the Internet as an alternative information source to learn about the COVID-19 pandemic. Misinformation had started flooding the web right from the initial stage of the emergence of the novel coronavirus, mainly from user-created content on social media [23]. Thus, as people turned to the web searching for information, there was limited, non-technical information for the non-expert public about the coronavirus. People rather got exposed to either learning incorrect information about SARS-CoV-2 and COVID-19 or embrace fake news, misinformation, and conspiracy theories, with grave consequences [16]. Some of the unfounded misinformation includes misconstruing COVID-19 as a "common cold", or as a hoax, which made people have a false sense of immunity, while others ignored any public health safety measures [11]. Similarly, the conspiracy theories propagated online include COVID-19 as a bioweapon, China virus that was intentionally released to reduce the world population, while 5G technology contributed to the fast spread of the pandemic. These beliefs initially led to the hoarding of essential goods, racial attacks against the Chinese and other Asians [14,15,24,25].

Other studies examined the role of social media and Internet news outlets in generating misinformation, disinformation, fake news, and conspiracy theories about COVID-19 [24,26,34,35]. These studies tend to leave out the aspect of "web search," such as the use of Google search, which constitutes a major channel through which the public obtain health-related information [28,29]. This paper undertakes the first empirical investigation using a "web search" to learn about SARS-CoV-2 and COVID-19 and peoples' attitudes toward public health guidelines as expressed in the following research objectives:

- determine what people learned about the COVID-19 pandemic through "web search."
- examine any association between what people learned about COVID-19 and behavior towards the public health guidelines.
- analyze the impact of misinformation and conspiracy theories about the COVID-19 pandemic on people's behavior towards public health measures.

These objectives are developed into research hypotheses in the Sections that follow.

Theoretical Background

The Connectivism Learning Theory

This section examines the connectivism learning theory, which explains using digital platforms to enable learning [30,31]. This study employs this approach to explore how people learned about SARS-CoV-2 and COVID-19 through 'web search' and the potential behavioral implications towards public health guidelines, which scientists and medical experts recommend as ways to check the spread of COVID-19. For example, the study investigates if learning through 'web search' helped people acquire accurate knowledge or misinformation and conspiracy theories about the COVID-19 pandemic and its implications. Also, recent studies show that many people are yet to understand the science and the concept of the novel coronavirus (SARS-CoV-2) and the disease, COVID-19, which increases the danger of embracing misinformation [14,34]. Several web platforms, including social media, online news, and other Internet channels contribute significantly to misinformation and conspiracy theories [14,32].

As proposed by George Siemens [31], the connective learning theory analyzes the use of digital devices, computer networks, and electronic platforms to learn. The view is considered a pedagogical strategy for the digital age, emphasizing knowledge sharing across an interconnected Web and Internet network [30,31]. The approach focuses on knowledge acquisition utilizing information technology platforms and learning from multiple sources, developing skills, and disseminating information [30]. The platforms incorporate information on social media, Internet websites or blogs, and search engines that users can employ to learn and exchange knowledge, skills, and expertise [31,34].

One of the implications of the connectivism learning theory is that learning can occur outside the traditional classroom to use networked systems that enhance connections, interactions, and collaborations among the learners [35]. However, some learning theory experts criticize the connectivism theory for not offering any improvement to the actual learning method other than utilizing the Web 2.0 and related platforms [30,31]. Hence it cannot be deemed a substantive learning theory. Instead, it provides a bridge to other pedagogical methods: behaviorism, cognitivism, and constructivism. The core of the Siemens and Downes connectivism idea aims to move away from the traditional classroom learning techniques to a new theory of learning that embraces technology as the learning tool, which can inspire the new generation of learners and educators [30,31]. Thus, the theory draws its strength from Web-based activities [35].

The key benefit of the method is its intuitiveness and the ability to captivate learners due to the ubiquitous use of the Internet in today's world. The following principles contribute to the popularity of connectivism as a learning theory [31].

i. Learning and knowledge rest in diversity of opinions, as experienced today.
ii. Learning is a process of connecting specialized nodes or information sources.
iii. Learning may reside in non-human appliances.
iv. The capacity to know more is more critical than what is currently known.
v. Nurturing and maintaining connections help to facilitate continual learning.
vi. Ability to see connections between fields, ideas, and concepts is a core skill.
vii. Currency (accurate, up-to-date knowledge) is the intent of all connectivism learning activities.
viii. Decision-making is itself a learning process. Choosing what to learn and the meaning of incoming information is seen through a shifting reality.

The connectivism learning theory, as explained above, closely mirrors the use of Google search trends and other Internet platforms to learn about the outbreak of SARS-CoV-2 and COVID-19, especially where the masses did not get adequate, timely information about the coronavirus from the public agencies [67,68].

The connectivism learning theory, which is well suited to personal study and self-regulated learning [36,37], in this case how individual members of the public learned about SARS-CoV-2 and COVID-19 using web search in the first six (6) months of the COVID-19 pandemic.

An Overview of the SARS-CoV-2 Outbreak and COVID-19 Pandemic

The Global Impacts of SARS-CoV-2 Outbreak

The COVID-19 pandemic has inflicted severe problems ranging from health crises to psychological, social, business, and economic consequences over the world [16,18,38]. Meanwhile, there is currently no specific cure for COVID-19. However, there has been significant progress and technological advances leading to substantial breakthroughs in vaccine discovery and development through the pioneer efforts by Pfizer, Moderna, and others from UK, India, China, and other countries [39]. Administering the COVID-19 vaccines is ongoing worldwide, while several other vaccine discoveries and developments are in progress [40]. In the meantime, ongoing prevention, monitoring, and public health awareness are essential to mitigate the public health and economic burden. The most important prevention strategy is to understand the disease and how it spreads.

Transmission

The coronavirus transmission is primarily through respiratory droplets released from infected persons during cough, sneeze, or speech. One can also contact the virus via contact with contaminated surfaces. The virus can remain infectious in the air for 3 hours and on inanimate surfaces for 2 to 3 days up to 9 days or longer. This has implications for nosocomial spreads and super-spreading events [41]. The virus has also been isolated from blood, urine, and stool specimens. It is important to note that asymptomatic infected people may not be aware that they are infected because they do not have the symptoms or may not recognize the symptoms. Infected individuals can be contagious for up to 4 weeks and can unknowingly be spreading the infection[41].

Clinical Presentation and Diagnosis

Symptoms usually appear 2 to 14 days after exposure. Most confirmed cases of SARS-CoV-2 infection are asymptomatic, and they recover without treatment. Common symptoms include fever, cough, shortness of breath, chills, myalgia, headache, sore throat, anosmia, and dysgeusia. Severe cases present with dyspnea, tachypnea, hypoxia (blood oxygen saturation </=93%), the arterial partial pressure of oxygen to fraction of inspired oxygen [PaO2/FiO2] less than 300, lung infiltration [27]. Some patients present with gastrointestinal

symptoms such as vomiting, diarrhea, abdominal pain, and cardiovascular features such as arrhythmia, shock, and acute cardiac injury [42]. There have been reports of asymptomatic carriers presenting with symptoms such as loss of smell and taste. In children, the majority present with mild (fever, cough, fatigue, congestion), moderate (pneumonia) symptoms [42]. Some may be asymptomatic. Children < 5 years old may present with respiratory organ failure.

Chest CT scan shows a distinct appearance of ground-glass lung opacity, often bilateral, in patients who develop pneumonia [27]. Other radiographic features such as "crazy-paving sign, multifocal organizing pneumonia, and architectural distortion in a peripheral distribution" may appear with disease progression. Diagnostic testing is performed from respiratory (nose, throat, saliva) and serum samples, using a real-time Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) panel or antibody test. The viral RNA has also been detected in stool and blood [5].

Complications

Some hospitalized patients develop thromboembolism, especially deep venous thrombosis, and pulmonary embolism. Other complications include Microvascular thrombosis of the toes, clotting of catheters, myocardial injury with ST-segment elevation, and large vessel strokes. This complication may be associated with the release of high levels of inflammatory cytokines and activation of the coagulation pathway caused by hypoxia and systemic inflammation secondary to COVID-19 [43].

Prevention and Control

People must be well-informed. Infected persons must practice respiratory etiquette to avoid infecting others, including covering coughs and sneezes with a tissue and discarding it properly, coughing into the inside of the elbow, and covering the nose and mouth properly with a surgical face mask. Best practices include proper handwashing with soap and water for at least 20 seconds or at least 60% alcohol-based hand rub. Clean touched surfaces with disinfectants frequently. Avoid touching the eyes, nose, and mouth with unwashed hands. Avoid close contact with people who are ill [44]. CDC recommends that infected and exposed individuals must isolate or quarantine themselves, respectively, for at least 14 days. The CDC also recommends social distancing (avoid mass gatherings or large community events, shaking hands, or giving "high fives") [40]. In healthcare settings, standard contact, and airborne precautions, as well as eye protection, should be used to mitigate the spread of SARS-CoV-2 [44]. There is no specific cure for COVID-19. Management is mainly supportive care and treatment of secondary infections. Severely ill patients may need advanced organ support.

Methods

Google Trends and Search Keywords about SARS-CoV-2 and COVID-19

This paper uses an infodemiology approach to evaluate the use of web search to learn about SARS-CoV-2 and COVID-19. As an area of science research, infodemiology is a method or technique designed to measure and track health information "demand" automatically (e.g., by analyzing search queries) as well as "supply" on the Internet [18,20,69]. The goal is to inform public health policy and practice. The study utilizes the Google trends data, a freely available online resource that provides information on what was/is trending based on actual users' Google queries [11,45].

Google trends offer various options, such as 'Trending Searches' (i.e., trending queries for daily search trends and real-time searches in a selected region) or 'Year in Search' (i.e., what was trending in a specific area in a particular year). Another option is to 'Explore' (which allows an investigation of an area of interest based on keywords over the selected periods and regions). This study utilizes the "Explore" option, which allowed data to be retrieved directly from the 'Google Trends Explore' page in .csv format. It is also important to note that Google Trend data points are normalized to have a maximum value of 100 and a minimum of zero (0). The normalization implies dividing each data point is by the total searches of the geography and time range it represents to compare relative popularity. Note that the value 0 does not necessarily indicate no searches but shows a significantly low search volume that does not warrant inclusion in the results [69].

In this study we captured the worldwide Google trends data covering the initial months of the SARS-

CoV-2 outbreak from 1st January to 30th June 2020 (i.e., 182 daily data points for each search term). Regarding the search terms, the paper employs twenty-five (25) keywords and phrases used by the public to learn about the COVID-19 pandemic through web searches. We identified the search keywords (Table 1) through a literature survey of published documents indexed on the Web of Science. Six (6) search terms relate directly to the ongoing pandemic (nCoV, 2019-nCoV, SARS-CoV-2, COVID-19, pandemic, coronavirus). Another six keywords address previous viral/coronavirus epidemics (e.g., SARS-CoV, SARS, MERS-CoV, MERS, Virus, Influenza). The third category of search terms represents public health safety measures that experts recommend as guidelines to limit the spread of COVID-19 (social distancing, wearing a facial mask, washing hands). The final category of keywords represents misinformation and conspiracy theories, such as "China Virus," "Common Cold," "Bioweapon" (Table 1).

Table 1. The learning terms about COVID-19 pandemic, misinformation and conspiracy theories, and the public health safety guidelines (based on a literature review).

Web Search/Learning Terms Used to Learn about SARS-CoV-2 and COVID-19	
COVID-19 & Related Epidemics	2019-nCoV [3]; “nCoV” [3,9]; “SARS-CoV-2” [41]; “COVID-19” [9,24,46]; “Pandemic” [6,7]; “MERS-CoV” [47,48]; “MERS” [47,49]; “SARS-CoV” [41,48]; “SARS” [48,49]; Virus [50]; “Coronavirus” [2,49,51]; “Influenza” [44]; FLU [61].
Misinformation / Conspiracy Theories	Virus Hoax [14,15]; “Injecting/Ingesting Bleach” [52]; “5G” technology enhancing the spread of the virus [24,26]; “COVID-19 Hoax” [14,15,25]; “Common Cold” (CommonCold2020) [11]; “China Virus” [46]; “Bioweapons” created by China [14,24]
Public Health Measures	“Social Distancing” [53]; “Wash Hands” or “Hands Wash” [53]; “Wear a Facial Mask" [54]; “Isolation” [53]; “Quarantine” [53].

Research Hypotheses

Hypothesis #1:

For the purpose of determining what people learned about the COVID-19 pandemic through "web search," we define the null and alternative hypotheses thus:

- $H1_0$: People did not learn about COVID-19 through ‘web search’ using the identified keywords
- $H1_1$: People learned about COVID-19 through ‘web search’ using the identified keywords

Hypothesis #2:

Based on the literature, using a "web search" to learn about a subject of interest can influence the learner's decision-making and actions [55]. On this premise, the study examines any association between what people learned about COVID-19 and people's behavior towards the public health guidelines. We develop two separate hypotheses. The first of hypotheses here relating to "web search" to learn about COVID-19 (concept, science, and structure of SARS-CoV-2 and COVID-19), while the second aspect evaluate learning about misinformation and conspiracy theories, and the behavioral response to the public health measures. The null and alternative hypotheses are as follows:

- $H2_0$: There is no association between what people learn about COVID-19 through "web search" and behavior towards public health measures.
- $H2_1$: There is an association between what people learn about COVID-19 through "web search" and behavior towards public health measures.

Hypothesis #3:

There is a widely held assertion that misinformation and conspiracy theories about the COVID-19 pandemic have had a significant impact on people's behavior towards the public health measures.

- $H3A_0$: There is no association between misinformation learned about COVID-19 and peoples' behavior towards public health measures.
- $H3A_1$: There is an association between misinformation learned about COVID-19 and peoples' behavior towards public health measures.

Similarly, we defined the null and alternative hypotheses for learning about conspiracy theories thus:

- $H3B_0$: There is no association between conspiracy theories learned about COVID-19 and peoples' behavior towards public health measures.
- $H3B_1$: There is an association between conspiracy theories learned about COVID-19 and peoples' behavior towards public health measures.

Data Analysis

Data analysis employed statistical trends and graphical visualization, correlation and regression, principal component analysis (PCA), and predictive models (12,52,63-65). The statistical trends and analyses involves evaluating relationships among the listed variables using the statistical trends, including graphical display, correlation, and principal components analysis (PCA), which helps determine the predictiveness of the learning attributes and learners' actions towards public health guidelines. We use the statistical analysis package called JMP, one of the SAS statistical software packages [57], and Microsoft Excel to create the charts and graphs. We also utilize SPSS to compute the correlation matrix and the PCA and statistical packages in R for the linear modeling. The evaluation helps establish the correlation between the study attributes.

Results

The data analyzed in this study came from the 'Google Trends' worldwide index covering the period from the initial outbreak of the novel coronavirus on January 1, 2020, up to June 30, 2020, when the pandemic became widely known (Huang et al. 2019). The outbreak was formerly reported to the WHO's office in China, on December 31, 2019 [6]. The reason for focusing on the first six (6) months of the pandemic is to capture what people learned during the early days after the outbreak and the possible impacts of what people learned through "web search" on individual's actions towards public health safety measures.

To better understand the characteristics of Google trends data, we present the summary statistics of the daily search index for each of the twenty-five (25) keywords or search terms/phrases (Table 2 in the Multimedia Appendix). The average normalized scores for the terms vary from 2.65 (Ingesting Bleach) to 39.75 (SARS-CoV-2) as shown in the said Table 2.

Temporal Trends: Using Web Search to Learn about SARS-CoV-2

The keywords employed to conduct web searches indicate what people learned about the COVID-19 pandemic [56]. As presented in Table 1, some of the search keywords addressed the novel coronavirus directly, while others examined misinformation and conspiracy theories.

Figure 1 presents the first category of web search terms that people utilized to learn about the COVID-19 pandemic. In the early period, most people used keywords and phrases that explain previous coronavirus epidemics, including "Influenza," "MERS," "MERS-CoV," "SARS," and "virus." Although scientists ruled out the past epidemics, the WHO officials highlighted those terms as examples of past coronavirus outbreaks during press briefings [6]. The use of those keywords in the web search nosedived after WHO formerly named the novel coronavirus and the disease ("nCoV," "2019-nCoV," "SARS-CoV-2," and "COVID-19"). The coefficient of determination (R^2) of the keywords are (SARS-CoV-2: 0.37; COVID-19: 0.36; Influenza: 0.27; 2019-nCoV: 0.24; nCoV: 0.18; SARS: 0.12; SARS-CoV: 0.11), indicating the proportion of the variation in the search index over the period for the listed keywords. Similarly, the following search terms (MERS, MERS-CoV, Pandemic, Virus, Coronavirus, nCoV, and 2019-nCoV) had (R^2 < 0.1).

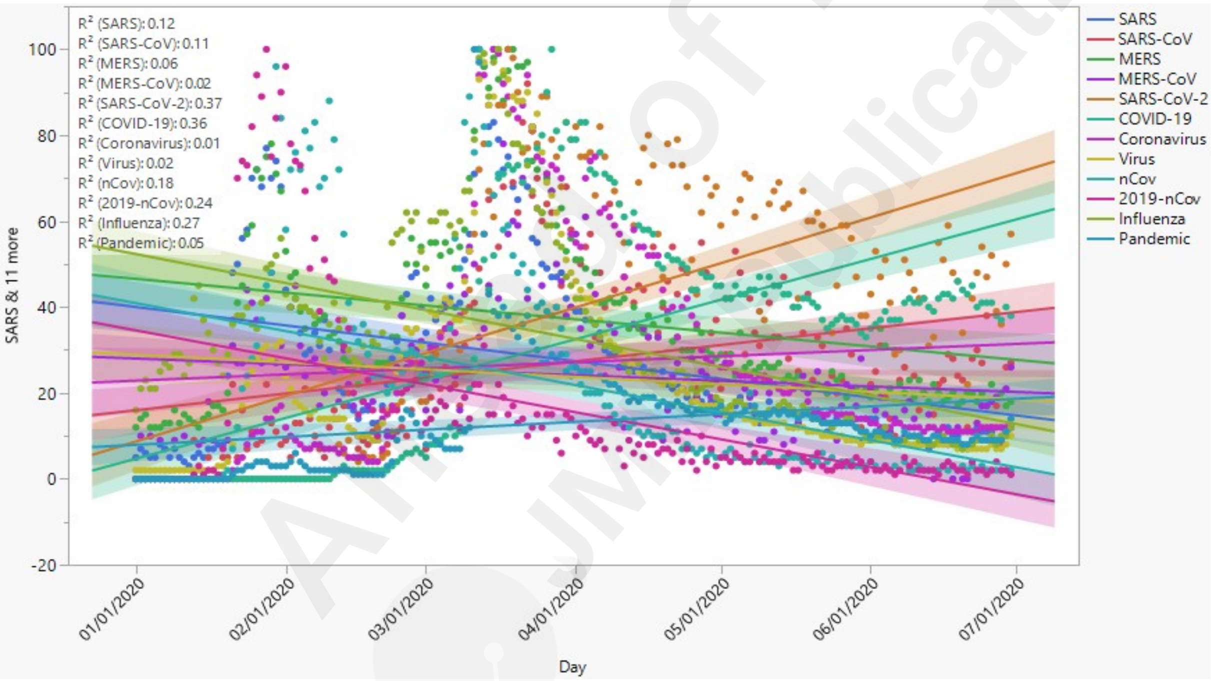


Figure 1: The keywords used by people to learn about SARS-CoV-2 and COVID-19 through we search.

The second category of the keywords involved misinformation and conspiracy theories (Figure 2). The variation in the use of the terms using the coefficient of determination (R^2) was "ChinaVirus:" (0.15), "Common Cold:" (0.10), and "5G:" (0.09). Most searches in the initial months of the outbreak utilized the keywords "Common Cold" (cold2020), "Biological Weapon," and "China virus." Thus, the misconception about SARS-CoV-2 as "common cold," a "biological weapon," or "China Virus" [46]. Some studies explain the purpose of releasing the coronavirus to include reducing the world population [14,24]. However, web searches using these terms fell continuously over time to a near-zero search index, while new words ("5G" and "COVID-19 hoax") surfaced and increased significantly (Figure 2).

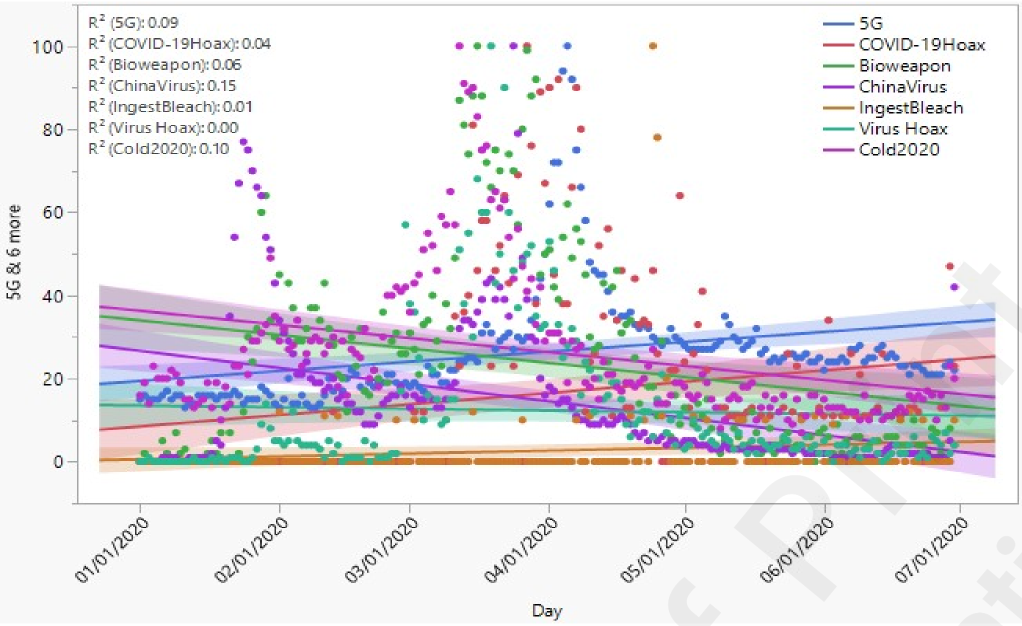


Figure 2. Monthly worldwide search index showing the learning terms that represent misinformation and conspiracy theories about COVID-19.

The third segment of the trend analysis involves the web search to learn about the public health measures (Figure 3). The results show that there was little or no interest in learning about wearing a facial mask ("wear a mask") and maintaining social distancing ("Social Dist") at the start of the pandemic. But the trends changed quite quickly, recording a dramatic increase from zero at the beginning of the outbreak to achieving a maximum search index in the months of March and April 2020, as the pandemic spread worldwide. The coefficients of determination were as follows: "Wear Mask" (R^2=0.56), social distancing (R^2=0.13), and Quarantine (R^2=0.09). The increases, especially regarding "facial masks" and "social distancing," were sustained for a long time.

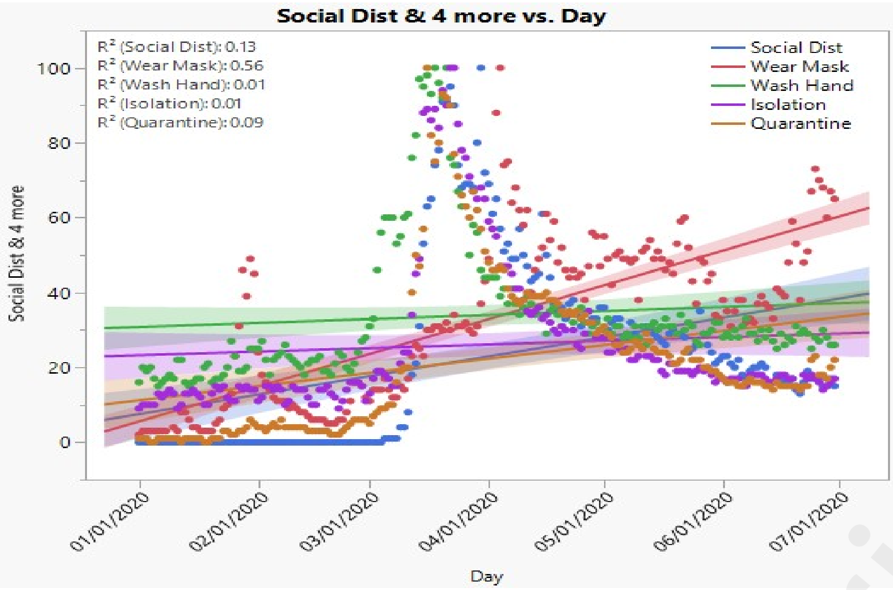


Figure 3: The trend analysis showing the public's interest in learing about public health safety measures

What Did People Learn about COVID-19 through Web Search?

The keywords identified above approximate what people learned about SARS-CoV-2 and COVID-19 through web search. However, some of the search keywords pre-existed the ongoing pandemic, while some terms refer to previous coronavirus epidemics (e.g., SARS, SARS-CoV, MERS and MERS-CoV, Influenza, Virus, Pandemic, and Coronavirus). It is plausible to argue that the search index for the listed pre-existed keywords represent purposes other than learning about COVID-19. Based on this assumption, we conduct a dependent two-sample t-test to examine the difference in the means search index of the pre-existed keywords in the previous years before COVID-19 outbreak and during the ongoing pandemic. Figure 4 compares the mean search index before and after the outbreak for each keyword.

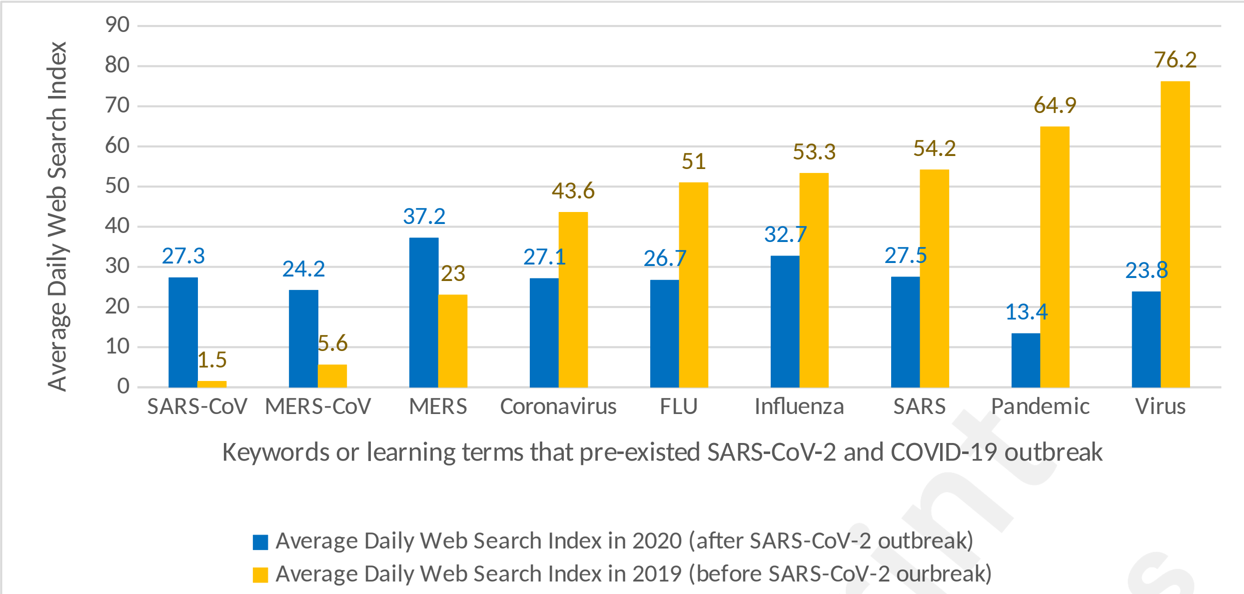


Figure 4. The average search index for pre-existed keywords before and during COVID-19 pandemic.

The null hypothesis is that the search indexes of each learning term (SARS and SARS-CoV, MERS and MERS-CoV, Influenza, Virus, Pandemic, and Coronavirus) before and after the outbreak of the ongoing COVID-19 are equal, and the alternative hypothesis is that they are unequal. The results (Figure 4) show that the differences in mean search index before and during the ongoing pandemic > 60% in all cases. Also, the p-values close to zero (0) for all the variables. We reject the null hypothesis and conclude that the significant differences in the mean search index of the variables were due to the ongoing COVID-19 pandemic. Unexpectedly, the mean search index for some pre-existed variables (e.g., FLU, Influenza, SARS, Pandemic, and Virus) declined during the pandemic.

What Search Term Contributed to Learning about SARS-CoV-2 and COVID-19?

The principal components analysis (PCA) was employed to evaluate the underlying latent variable of the search terms that contributed to learning about COVID-19. Based on the scree plot and the elbow rule, we can limit the factors extracted to the first two principal components (Figure 5): the keywords that address previous epidemics (pastCoVepidemics) and the terms that explain the ongoing pandemic (presCoVpandemic). Scree plot is a graphical representation of the percentage variability explained by each principal component.

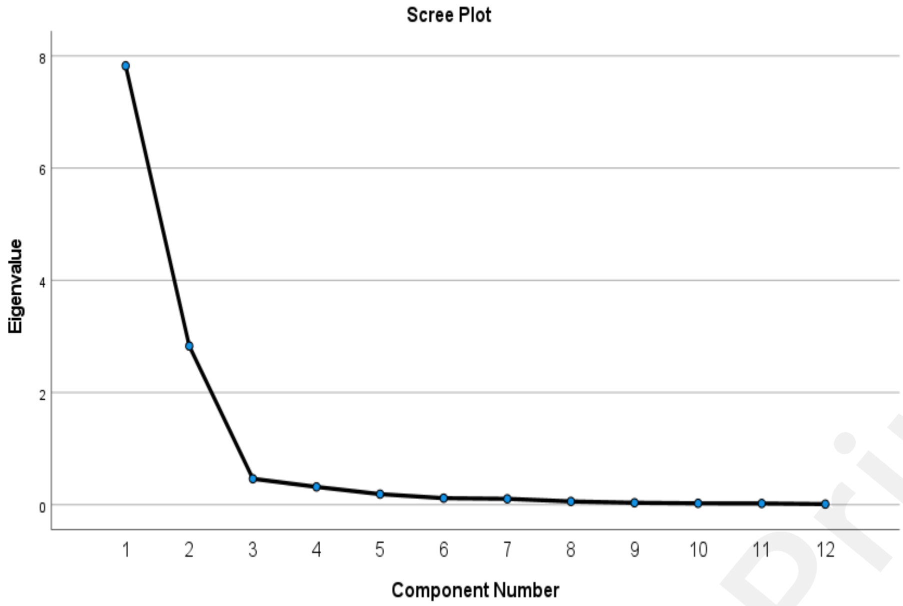


Figure 5: Scree plot of COVID-19 learning terms; two principal components extract 88.78% of total variation

The first two underlying components explained a total of 88.78% of the variation in learning terms, with the first component (pastCoVedpidemics) and second (presCoVpandemic) determining about 65.2% and 23.58%, respectively, of the information about COVID-19 from the twelve search keywords (Table 3a).

Table 3a Total explained variance for search keywords used to learn about COVID-19

Component		PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12
Initial Eigen-values	Total	7.82	2.83	0.46	0.32	0.19	0.12	0.11	0.06	0.04	0.03	0.02	0.01
	% Var	65.20	23.58	3.86	2.65	1.57	0.98	0.88	0.49	0.29	0.21	0.19	0.09
	Cum %	65.20	88.78	92.64	95.30	96.87	97.84	98.72	99.21	99.51	99.72	99.91	100

A linear combination of the two components (pastCoVepidemics and presCoVpandemic) are:

- PastCoVepidemic = 0.98 Virus + 0.934 Corona virus + 0.929 MERS + 0.923 Flu + 0.858 MERS-CoV + 0.858 SARS + 0.791 SARSCoV + 0.799 Pandemic + 0.814 Influenza
- PresCoVpandemic = -765 nCov + 0.784 COVID19 + 0.766 SARSCoV2

Table 2b shows the weights (loadings) of the terms for the two components. Note that we record loading greater the 0.6 to combine only search keywords that have a high correlation with the component in the linear combinations.

Table 2b Component matrix/weight loading for search keywords used to learn about COVID-19

		Virus 2020	Coronavirus2020	MERS 2020	Flu 2020	MERS CoV2020	SARS 2020	Influenza 2020	Pandemic 2020	SARS CoV2020	COVID19	SARS CoV2	2019 nCoVAV
Com-ponent	1	0.98	0.934	0.929	0.923	0.858	0.858	0.814	0.799	0.791	0.562	0.601	0.45
	2	-0.059	0.278	-0.305	-0.278	-0.121	-0.443	-0.477	0.315	0.505	0.784	0.766	-0.765

What Terms Fueled Misinformation and Conspiracy Theories about COVID-19?

We identified eight (8) search keywords from the literature that denote misinformation and conspiracy theories (Table 1). We also perform a PCA to evaluate the search terms that fuelled misinformation and conspiracy theory. The results (Table 4a; Figure 6) identify three (3) principal components and the variability (PC1: 48.17%; PC2: 22.65%, and PC3: 14.03%) based on the elbow rule.

Table 4a Total variance explained involving terms that fueled misinformation and conspiracy theory

Component		PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
Initial Eigen-values	Total	3.372	1.586	0.982	0.496	0.257	0.176	0.132
	% of Var	48.171	22.652	14.026	7.085	3.669	2.509	1.888
	Cum%	48.171	70.823	84.849	91.934	95.603	98.112	100

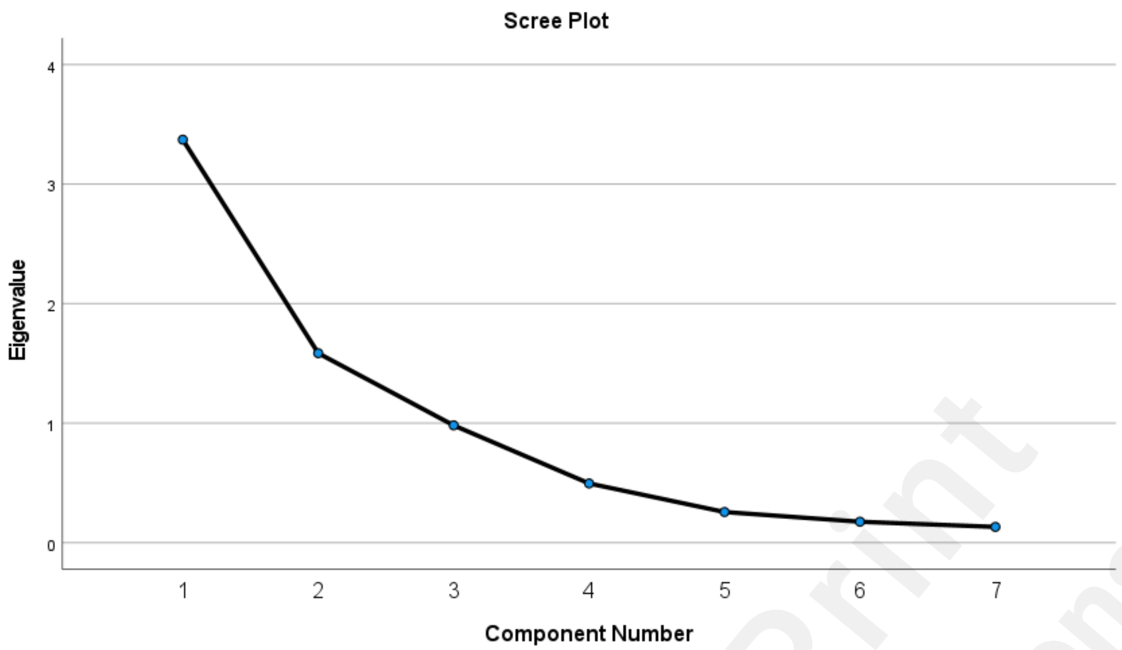


Figure 6: Screep plot of misinformation and conspiracy theory terms

Table 4b Component matrix/weight loading for terms that fueled misinformation and conspiracy theory

		Bioweapon 2020	VirusHoax	CommonCold 2020	COVID-19Hoax	ChinaVirus	@5G	IngestBleach
Components	1	0.928	0.908	0.789	0.692	0.601	0.471	-0.045
	2	-0.152	0.033	-0.438	0.624	-0.558	0.786	0.226
	3	0.03	0.005	0.093	-0.044	0.059	-0.145	0.973

The three (3) components explain 84.85% of the variation in the search keywords under misinformation and conspiracy theory. The first component represents misinformation. We can quantify the daily number of misinformation searched using the linear combination.

- Misinformation=0.789 Common Cold + 0.928 Bioweapon + 0.908 Virus Hoax + 0.875 Cold 2020 + 0.692 COVID19Hoax + 0.60 China Virus.

The second and third components address the conspiracy theories (ConspTheory1, ConspTheory2) and the speculation that “5G” technology contributes to spreading COVID-19; “COVID-19 is ChinaVirus” intentionally created/released,” and the assertion that “ingesting or injecting bleach” can cure COVID-19 infection/kill the virus.” The results present this variable in a separate component (Table 4b).

- ConspTheory1 = 0.786 @5G

- ConspTheory2 = 0.97 IngestBleach

Public Health Safety Measures

This section investigates the impacts of what people learned through web search on behaviors towards public health safety measures against COVID-19. Based on the elbow rule, the PCA identifies two components, which we label as “PubHealthMes1” and “PubHealthMes2.” The two components account for the variability in the ‘Search Index’ of keywords used to learn about the public health measures against the spread of COVID-19 (Figure 7). The first and second components explain more than 75.4% and 18.7% of the variability respectively. That is the first two component explains about 94% of the variability (Table 5a).

Table 5a Total explained variance for terms that explain the public health measures against COVID-19

Component		1	2	3	4	5
Initial Eigenvalues	Total	3.768	0.933	0.24	0.035	0.023
	% of Var	75.359	18.665	4.799	0.71	0.466
	Cum%	75.359	94.025	98.824	99.534	100

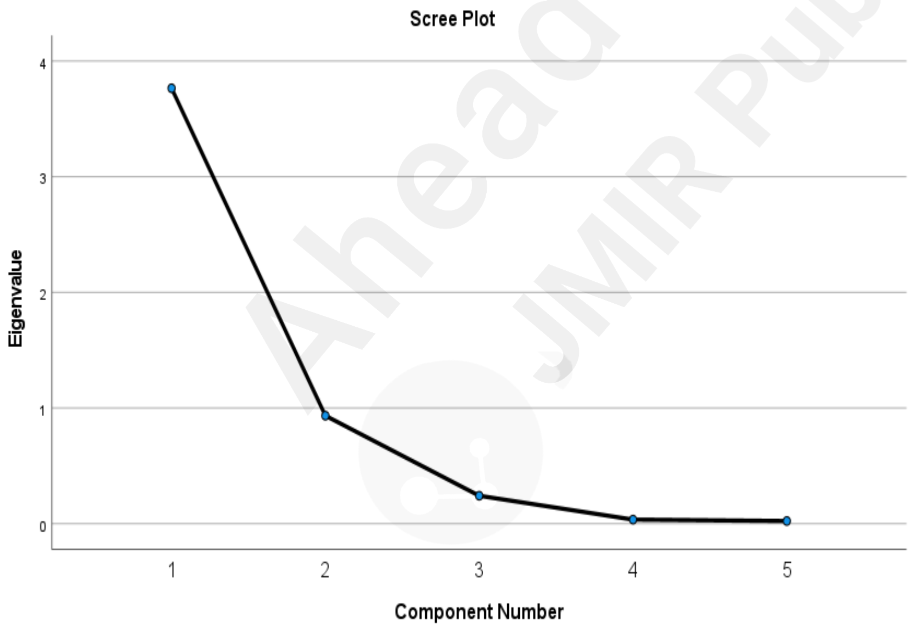


Figure 7. Scree plot of public health measures against COVID-19; two component extract 93.99% of the variability in the search index

The first component, “PubHealthMes1,” involves “social distancing,” “washing hand,” “isolation,” and “quarantine.” The second component, “PubHealthMes2,” with “wear mask” alone explains 84.7% variability.

- PubHealthMes1 = 0.953 Social Distance +0.847 Wash Hand + 0.953 Isolation + 0.99 Quarantine
- PubHealthMes2 = 0.847 WearMask

Table 5b Component matrix/weight loading for terms that explain the public health measures

		Quarantine	Social Distancing	Isolation	Wash hand	Wear Mask
Components	1	0.99	0.953	0.953	0.847	0.503
	2	-0.028	0.143	-0.223	-0.38	0.847

Analysis of the Relationships Among the Principal Components

This Section presents further analysis to test the hypothesis raised in the earlier Section using built predictive models. The variables identify the linear combination of significantly correlated search keywords (loading >0.6) to the principal components discussed in the results above. Also, Tables 3a/3b, 4a/4b, 5a/5b present the underlying latent variables of the twenty-five (25) search terms used to learn about CIVID-19, the misinformation and conspiracy theories, and the public health measures. As stated, Figure 8 shows the underlying variables.

Here, we examine how the underlying variables and the search terms impacted learning and behavior towards the public health measures (learning about COVID-19: PastCoVepidemics and PresCoVpandemic; misinformation/unproven or misleading assertions: Misinformation; Conspiracy Theories: ConspTheory1 and ConspTheory2; and the public health safety measures: PubHealthMes1 and PubHealthMes2).

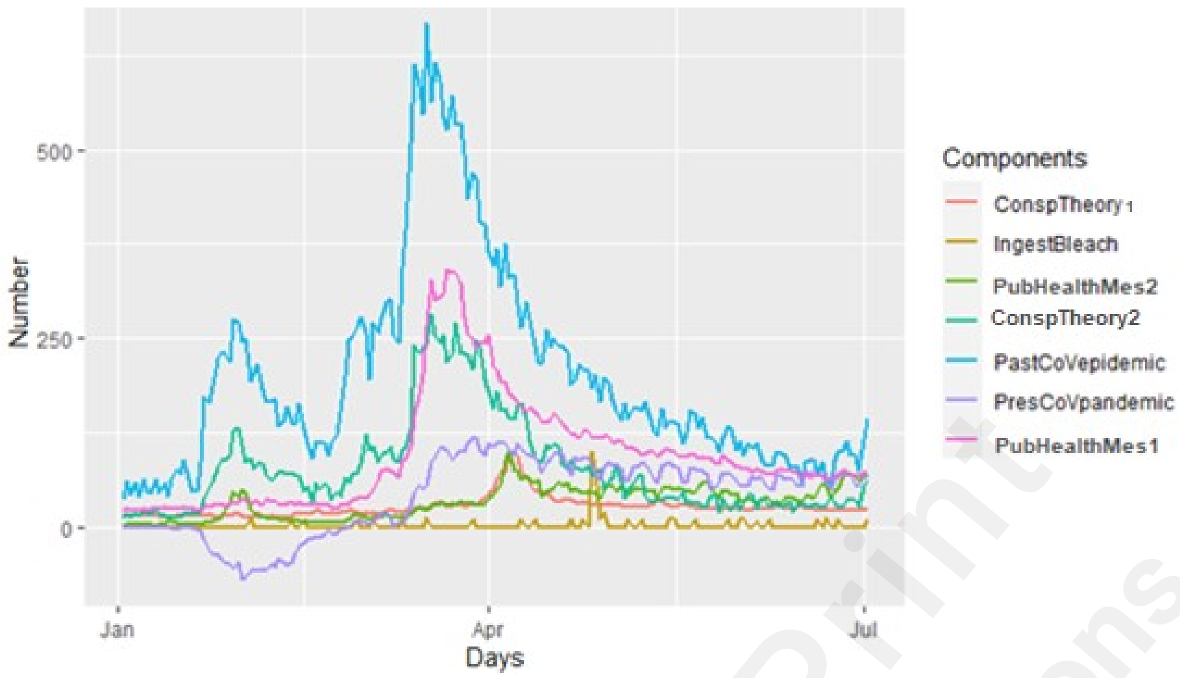


Figure 8. Daily search index of the principal components

The results show three essential highlights from the daily search index. First, the most popular search terms used at the initial outbreak of the pandemic in early January 2020 were terms representing the misinformation and the past epidemics. The terms that dominate web search during the early days of the outbreak represent previous coronavirus epidemics and the misinformation. The search keywords that represent the conspiracy theories were not used until May 2020. Also, the use of learning terms that directly explains COVID-19 (PresCoVpandemic) corresponds with WHO's naming and renaming of the coronavirus and the diseases (2019-nCoV, SARS-CoV-2 and COVID-19) in Jan, Feb, and Mar 2020.

Information Learned Versus Behavior Towards Public Health Measures

We employed correlation analysis among the variables, the scatter plot, and their histograms to examine the relationship between what people learned and attitude toward public health measures (Figure 9).

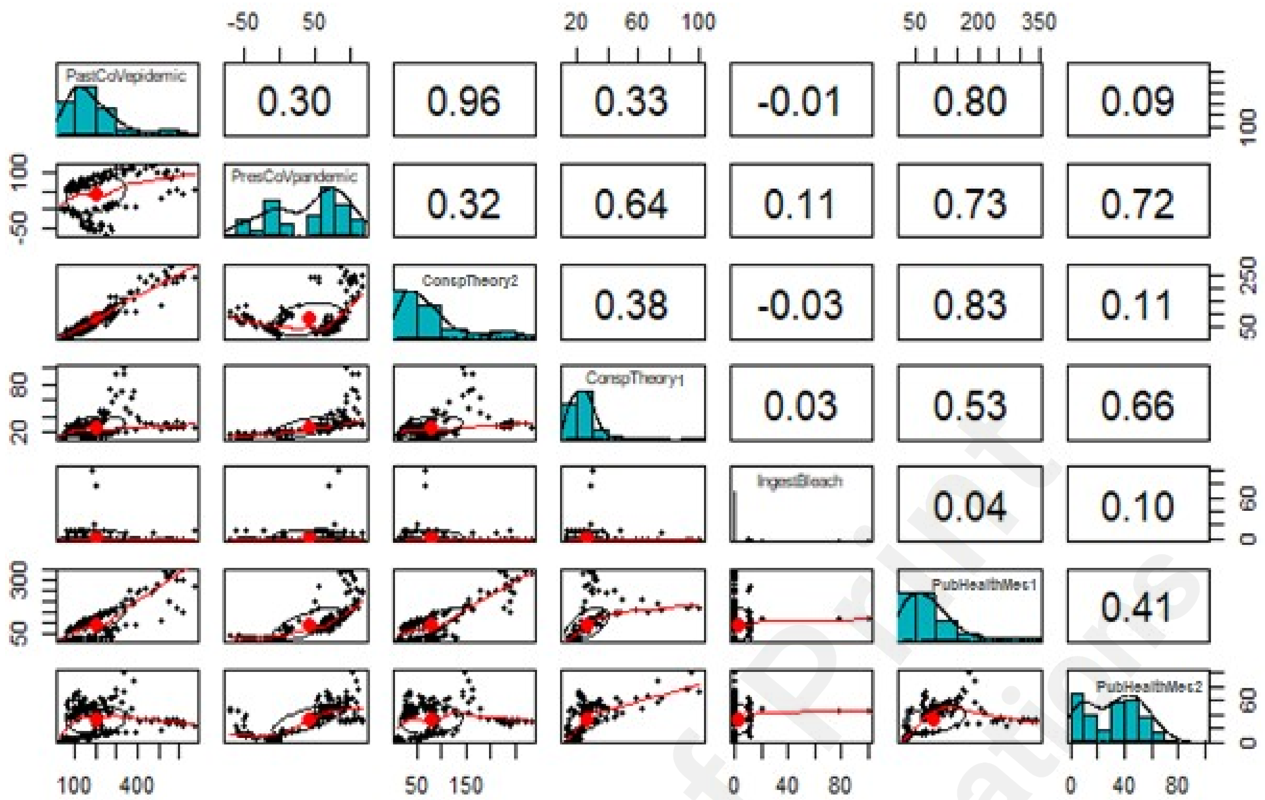


Figure 9. Panel Pair Plot of linear correlation

PubHealthMes1 has a robust positive relationship with pastCoVepidemics (r=0.80) and a moderate positive relationship with presCoVpandemic (r=0.73), which implies the effectiveness of learning keywords associated with past coronavirus epidemics (e.g., SARS, SARS-CoV, MERS, MERS-CoV, and more) and the ongoing pandemic (2019-nCoV, SARS-CoV-2, and COVID-19). Similarly, PubHealthMes2 (wearing a facial mask) has a moderate positive relationship with presCoVpandemic (r=0.71). There is a strong association between actions taken and the information learned. Figure 9 shows a correlation matrix.

Given such a strong linear relationship between the search terms and people's behavior and action, a multiple linear regression model seems acceptable as a predictive model. But the data fail the assumption of normality, as shown by the Q-Q plots in Figure 10a/b. A normal QQ plot helps to compare two probability distributions (plotting the residuals against theoretical quantiles). Figure 10 a/b indicates that most residuals are not lying on the diagonal line; hence the data is not normally distributed.

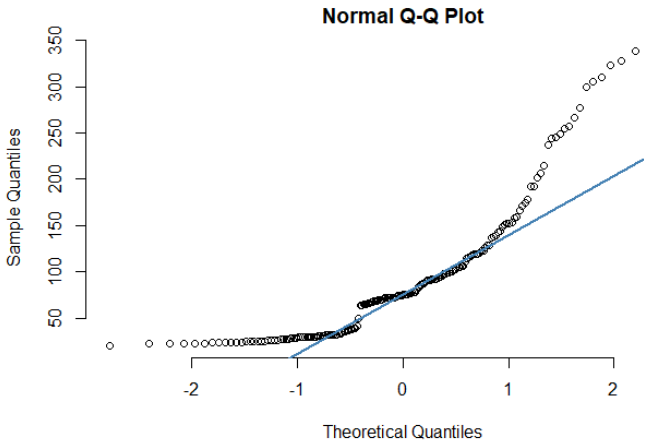


Figure 10a: Q-Q plot of PubHealthMes1

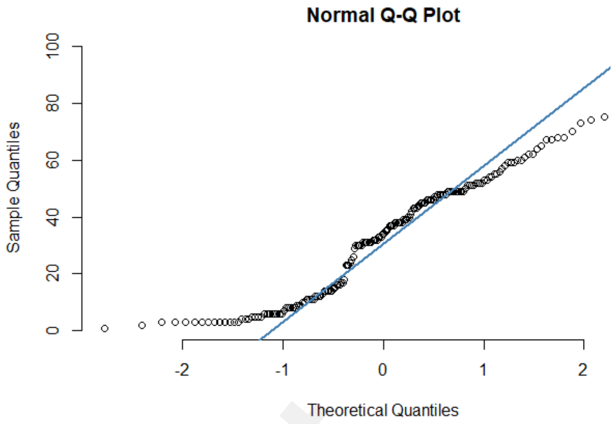


Figure 10b: Q-Q plot of the PubHealthMes2

From Figure 9, we can observe that the "PubHealthMes1" histogram suggests a right-skewed distribution, while the "PubHealthMes2" histogram suggests a bimodal distribution. To describe and explain the relationship between the public health measures and the terms that contributed to learning about COVID-19, misinformation, and conspiracy theories, we use the log linear predictive model since the data is based on a number of occurrences or frequency and not normally distributed. Also, the log linear model does not need to satisfy any assumptions, which we represent as follows:

$$\log(Y_i) = \beta_0 + \sum_{i=1}^{m} \beta_i x_i \; i = 1, \ldots, m$$

where x_i are the covariate and β_i are the parameters to be estimated.

The data obtained through the PCA have been fitted to a log linear regression model using Rstudio tools. The regression model obtained becomes:

Log(PubHealthMes1) = 3.33 + 0.0024 PastCoVepidemics + 0.011 PresCoVpandemic

The R^2 of the model is 0.93 and the p-value is less that 2.210^{-16}. All the coefficients in this model are significant at 0.0 alpha level. From the model, we see that there is 1.10 percentage (100(exp(0.011)-1)%) increase in the Public health measure for every one (1%) percent increase in PresCoVpndemic (2019-nCoV, SARS-CoV-2 and COVID-19). This may be explained by links contained present COVID pandemic articles on public health measures. For every one (1%) percent increase in the past epidemic search, there is a 0.25% increase in the search of the public health measures. Similarly, the regression model for PubHealthMes2" obtained becomes:

$$\text{Log(PubHealthMes2)} = 0.01\text{PresCoVpandemic} + 2.35$$

The R^2 of the new model is 0.50 and the p-value is less that 2.210^{-16}. The two models show that both PubHealthMes1 and PubHealthMes2 are significantly predictive using the PresCoVpandemic. It shows that learning about the present pandemic creates an incentive for people to learn about public health measures, which we can approximate to a desire intention to comply with the measures.

Misinformation, Conspiracy Theory, and Public Health Measures

This section examines the relationship between conspiracy theories, misinformation, and public health measures using correlation and predictive analyses. As discussed earlier, we recategorized the search terms representing misinformation and conspiracy theory have into two principal components. Hypothesis H3A (defined in the previous section) focuses on misinformation and public health guidelines. At the same time, H3B addresses the impact of conspiracy theories on people's behavior towards the same safety measures.

Using the correlation analysis to evaluate the association between the conspiracy theory and public health measures, we observed a moderate positive linear association between ConspTheory2 and PubHealthMes2 (r=0.66) and a moderate linear relationship between ConspTheory1 and PubHealthMes1 (r= 0.53). The "log-linear" analysis showed that conspiracy theory is not a significant predictor of PubHealthMes1 (p=0.622) but is for PubHealthMes2 (p=0.0084). Thus, the null hypothesis that there is no association between conspiracy theory (ConspTheory1) and people's response to public health (wearing facial masks to limit the spread of COVID-19) is rejected. We can conclude that conspiracy theory is predictive of people's behavior in wearing a facial mask.

We also analyzed the relationship between misinformation and behavior towards public health measures. Although the correlation and the predictive analyses show is a moderate positive linear relationship between misinformation and PubHealthMes1 (r=0.83), it is not predictive. The relationship between the two variables is just mathematical but not causal. Despite a negligible negative linear relationship between misinformation and wearing a facial mask (r=0.11), the "log linear" model shows that misinformation is not a significant predictor for both PubHealthMes1 (p =0.267) and PubHealthMes2 (p=0.13). Notwithstanding the strong linear relationship between web search to learn about misinformation and public health measures, there is no sufficient evidence to reject the null hypothesis. We can conclude that there is no association between learning about COVID-19 misinformation and people's response to the public health guidelines.

Discussion

Principal Findings: We found that people used search keywords related to past coronavirus epidemics (pastCoVepidemic) and the ongoing pandemic (presCoVpandemic) to learn about SARS-CoV-2 and COVID-19. However, the attention accorded to the pandemic led to less focus on the terms relating the perennial illness (e.g., common cold, flu, and more). These results corroborate studies reporting the unintended positive consequences of COVID-19 leading to declines in cases of "Influenza," "FLU," and similar infections [e.g., 44]. Other learning terms employed are keywords that address the pandemic directly. The average search index for those keywords include nCoV/2019-nCoV (19.01), SARS-CoV-2 (39.75), and COVID-19 (32.40).

Studies examining learning by "web search" emphasize the significance of the search terms or phrases to what the users intend to learn. A trending word on the web indicates what information people are interested in learning [11,55,56]. This study identified twenty-five (25) most utilized keywords to learn about SARS-CoV-2 and COVID-19 through the web search.

Regarding the impacts of what people learned on behavior towards public health measures, the PCA identified three (3) latent variables classified as Misinformation, ConspTheory1, and ConspTheory2. Only ConspTheory1(@5G) directly and significantly influenced people's behavior towards public health measures [PubHealthMes2(Wear Mask)]. The conspiracy that 5G technology enhances the easy spread of COVID-19 [14] highlights danger, which can cause people to take precautions. A different study [15] identified erroneous belief in the 5G conspiracy theory as leading to the hoarding of essential goods during the initial period of the SAR-CoV-2 outbreak. Although there was a high correlation between misinformation (Table 4a/

4b) and behavior towards public health measures, this was not statistically significant based on the web search index. Also, as the pandemic lingers, causing severe health and social crisis, strains in family relations, economic and business losses, many people become increasingly aware of the COVID-19 dangers [21,53,58-60]. Through direct impacts or by experience, this can cause people's change in behavior irrespective of whether they believed the misinformation or not.

Strengths and Limitations

Internet platforms continue to play a significant role in health communication during the ongoing COVID-19 pandemic. Some studies attribute the increase in misinformation and conspiracy theories about COVID-19 in different countries to web search, social media use, and online news media platforms to learn about SARS-CoV-2 and COVID-19 [10,11,35]. However, most studies are anecdotal with no empirical evidence. Using the Google trend data, this study provides the first empirical evidence to this discourse. In the era of big data, the analysis of Google queries can be envisioned as a valuable tool for researchers to explore and predict human behavior, especially as studies suggest that online data can correlate with actual health data (70,71).

Infodemiology studies have their limitations too. While Google search keywords are short and easy to classify automatically, interpreting the terms semantically can be challenging. It is not clear why people are searching for these keywords. Furthermore, when using Google Trends, the sample is unknown and may not be presentative, and individuals using the Internet are not representative of the entire population. They are more likely to be younger, more educated, earn higher incomes, and reside in urban areas [18]. Individuals who are more likely to be severely affected by COVID-19 are not usually represented by this population (72,73). Despite the identified limitations, previous studies suggest that Web-based data provide valuable and valid results in exploring and predicting behavior and highly correlate with actual data. [70,71]. Further, there are reports of rapid penetration of internet access and usage in different parts of the world except for regions with low internet penetration or countries with low scorings in freedom of speech [33,74,75].

Conclusion

The results of this empirical infodemiology study show that a good portion of the global population learned about the outbreak of SARS-CoV-2 and COVID-19 through a web search, particularly in the early period of the pandemic. The period covers the initial days, weeks, and months from the emergence of the novel coronavirus in Jan 2020 up to June 30, 2020, when the public became more aware of the pandemic, especially after the first wave [1].

The principal components analysis shows that people used the web to learn about the ongoing COVID-19 in two aspects, namely, using search terms relating to the past coronavirus epidemics (pastCoVepidemics) and utilizing keywords that directly address the ongoing COVID-19 pandemic (presCoVpandemic). The use of pastCOVepidemics keywords in the web search nosedived as WHO formerly named the novel coronavirus and the disease ("nCoV," "2019-nCoV," "SARS-CoV-2," and "COVID-19"), and the terms became available. The trends analysis shows that web searches to learn about COVID-19 followed a similar trend as learning about the public health measures, implying that the more people-focused attention on learning about SARS-CoV-2 and COVID-19, the more they also studied about the public health measures, and vice versa. Interestingly, learning about the conspiracy theory (ConspTheory1) that 5G technology contributes to the fast spread of COVID-19 globally is a predictor of people's behavior towards public health measures (PubHealthMes2). This erroneous belief makes people take precautionary measures of wearing a facial mask, although borne out of fear [14,15]. The same studies using the survey method also identify the same conspiracy theory (5G) as making people respond out of fear to taking precautions. This factor contributed to stockpiling of goods in the early days of the pandemic [15]. This study is the first to examine what people learned through web searches and how these influence people's social behavior towards the public health safety guidelines.

Conflicts of Interest

None Declared.

Multimedia Appendix 1

Google Trend Data Summary Statistics

Table 2: Summary statistics of the normalized daily global Google Trend scores for different keywords used in this study. Data correspond to the time window between 1st Jan 2020 to 30th June 2020 (n=182)

	Keywords	Min	Q1	Median	Mean	sd	Q3	Max
COVID-19 & Related Epidemics	2019-nCoV	0	3	7	15.62	22.43	16	100
	nCoV	0	3	10	21.93	25.75	31.75	100
	SARS-CoV-2	0	5	42.5	39.75	29.59	64.75	100
	COVID-19	0	2	39	32.39	26.8	47	100
	Pandemic	0	3	10.5	13.44	13.91	17	100
	MERS-CoV	0	12	20.5	24.16	17.25	31.75	100
	MERS	12	21	30	37.2	21.89	48.75	100
	SARS-CoV	0	12	23.5	27.34	20.17	40.75	100
	SARS	3	11	18.5	27.48	21.34	39.75	100
	Virus	2	8.25	14	23.76	23.56	30.75	100
	Coronavirus	0	11	16	27.13	25.97	38.75	100
	Influenza	7	16	27	32.68	21.92	42	100
	FLU	7	15	21	26.71	18.47	32	100
Misinformation	Virus Hoax	0	2	5	12.32	17.78	15	100
	Ingesting Bleach	0	0	0	2.65	10.04	0	100
	"5G" technology enhancing the spread of the virus	13	18	24	26.43	14.02	29	100
	COVID-19 Hoax	0	0	11	16.5	23.3	25.5	100
	CommonCold2020	7	15	19	26.38	17.87	30	100
	China Virus	0	2	9	14.59	18.22	18	100
	Bioweapons	0	4	14.5	23.79	24.46	34	100
	Social Distancing	0	0	18.5	22.85	24.45	33	100
	Hands Was	15	23	29	33.97	18.46	34	100

Public Health Measures	Wear a Facial Mask	1	12	34	32.73	21.2	49	100
	Isolation	9	14	18	26.1	20.72	28.75	100
	Quarantine	0	4	17	22.25	21.45	32.75	100

References

1. Taylor, A. Wuhan: The Chinese mega-city at the center of coronavirus outbreak. The Washington Post. January 23, 2020. https://www.washingtonpost.com/world/2020/01/23/wuhan-chinese-mega-city-center-coronavirus-outbreak/. [Accessed: Oct 31, 2020].
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z. Clinical features of Patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 2020 February 15-21;395(10223):497-506. doi:10.1016/S0140-6736(20)30183-5
3. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet. 2020 February 29-March 6;*395*(10225):689-697. doi:10.1016/S0140-6736(20)30260-9
4. Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S, Yuen KY. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging microbes & infections. 2020 Jan 1;9(1):221-236. doi:10.1080/22221751.2020.1719902
5. Woo J, Lee EY, Lee M, Kim T, Cho YE. An in vivo cell-based assay for investigating the specific interaction between the SARS-CoV N-protein and its viral RNA packaging sequence. Biochemical and biophysical research communications. 2019; Dec 10; 520(3):499-506. doi:10.1016/j.bbrc.2019.09.115
6. World Health Organization. WHO statement regarding cluster of pneumonia cases in Wuhan, China. 2020: https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china [accessed Dec. 1, 2020].
7. Labrague L, Yboa B, McEnroe–Petitte D, Lobrino L, Brennan M. Disaster preparedness in Philippine nurses. Journal of Nursing Scholarship. 2015; Dec. 09;48(1):98–105. Doi:10.1111/jnu.12186.
8. Atzrodt CL, Maknojia I, McCarthy RD, Oldfield TM, Po J, Ta KT, Stepp HE, Clements TP. A Guide to COVID‐19: a global pandemic caused by the novel coronavirus SARS‐CoV‐2. The FEBS journal. 2020 Sep;287(17):3633-50. doi:10.1111/febs.15375.
9. World Health Organization. Events as they happen: rolling updates on coronavirus disease (COVID-19). 2020: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. [accessed October 31, 2020].
10. Dadaczynski K, Okan O, Messer M, Leung AY, Rosário R, Darlington E, Rathmann K. Digital Health Literacy and Web-Based Information-Seeking Behaviors of University Students in Germany During the COVID-19 Pandemic: Cross-sectional Survey Study. J Med Internet Res. 2021 Jan 15;23(1): e24097. doi:10.2196/24097.
11. Rovetta A, Bhagavathula AS. COVID-19-Related Web Search Behaviors and Infodemic Attitudes in Italy: Infodemiological Study. JMIR public health and surveillance. 2020;6(2). doi:10.2196/19374
12. Sulyok M, Ferenci T, Walker M. Google Trends Data and COVID‐19 in Europe: Correlations and model enhancement are European wide. Transboundary and Emerging Diseases. 2020. doi:10.1111/tbed.13887
13. Akpan IJ, Aguolu OG, Akpan AA. An empirical analysis of what people learned about COVID-19 through a web search and the impacts on misinformation and attitude towards public health safety guidelines. medRxiv. Preprint posted online February 26, 2021. doi:10.1101/2021.02.23.21252323
14. Miller JM. Do COVID-19 conspiracy theory beliefs form a monological belief system? Canadian Journal of Political Science. 2020; 53(2), 319-326. doi: 10.1017/S0008423920000517
15. Miller JM. Psychological, political, and situational factors combine to boost COVID-19 conspiracy, theory beliefs. Canadian Journal of Political Science. 2020; 53(2);327-334. doi:10.1017/S000842392000058X
16. Allington D, Duffy B, Wessely S, Dhavan N, Rubin J. Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. Psychological medicine, 2020; 1-7. doi:10.1017/S003329172000224X
17. Hornik R, Kikut A, Jesch E, Woko C, Siegel L, Kim K. Association of COVID-19 Misinformation with Face Mask Wearing and Social Distancing in a Nationally Representative US Sample. Health communication, 2021; 36(1), 6-14. doi:10.1080/10410236.2020.1847437

18. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res. 2009. Mar 27; 11(1):e11 PMID:19329408.
19. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med. 2011 May;40(5 Suppl 2):S154-8. PMID:21521589
20. Mavragani A. Infodemiology and Infoveillance: Scoping Review. J Med Internet Res 2020;22(4). doi:10.2196/16206
21. Akpan IJ, Soopramanien D, Kwak DH. Cutting-edge technologies for small business and innovation in the era of COVID-19 global health pandemic. Journal of Small Business & Entrepreneurship, 2020; 1-11. doi:10.1080/08276331.2020.1799294
22. Neely, S., Eldredge, C., & Sanders, R. (2021). Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study. J Med Internet Res. 23(6). doi:10.2196/16206
23. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. International Journal of Medical Informatics, 2021; 148. doi:10.1016/j.ijmedinf.2021.104375
24. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data. J Med Internet Res. 2020; 22(5). doi:10.2196/19458
25. Garrett L. COVID-19: the medium is the message. The Lancet. 2020. *395*(10228), 942-943. PMID:32171075
26. Fleming, N. Coronavirus misinformation, and how scientists can help to fight it. Nature. 2020; 583: 155. PMID: 32601491
27. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, . . . Xia L. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. Radiology. 2020; 200642. DOI:10.1148/radiol.2020200642
28. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, Murugiah K. The use of google trends in health care research: a systematic review. *PloS one*, 2014; *9*(10). Doi:10.1371/journal.pone.0109583
29. Chung M, Oden RP, Joyner BL, Sims A, Moon RY. Safe infant sleep recommendations on the Internet: let's Google it. The Journal of pediatrics. 2012; 161(6), 1080-4. doi:10.1016/j.jpeds.2012.06.004
30. Goldie JGS. Connectivism: A knowledge learning theory for the digital age? Medical teacher, 2016; 38(10), 1064-1069. doi:10.3109/0142159X.2016.1173661
31. Dunaway KM. "Connectivism: Learning theory and pedagogical practice for networked information landscapes," Reference Services Review. 2011;39(4),675-685. doi:10.1108/00907321111186686
32. Mian A, Khan S. Coronavirus: the spread of misinformation. BMC medicine, 2020; 18(1), 1-2. doi: 10.1186/s12916-020-01556-3
33. Akpan IJ, Ibidunni AS. Digitization and technological transformation of small business for sustainable development in the less developed and emerging economies: a research note and call for papers. Journal of Small Business & Entrepreneurship. 2021. May 6; 1-7. doi:10.1080/08276331.2021.1924505
34. Lee JJ, Kang KA, Wang MP, Zhao SZ, Wong JYH, O'Connor S, ... Shin S. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res. 2020. 22(11). doi:10.2196/22205
35. Kim HK, Ahn J, Atkinson L, Kahlor A. Effects of COVID-19 Misinformation on Information Seeking, Avoidance, and Processing: A Multicountry Comparative Study. Science Communication. 2020. 42(5), 586-615. doi:10.1177/1075547020959670
36. Schunk DH. Self-regulated learning: The educational legacy of Paul R. Pintrich. Educational psychologist. 2005. 40(2);85-94. doi:10.1207/s15326985ep4002_3
37. Akpan IJ. Formative Assessment as a Technique for Improving Student Engagement and Learning Effectiveness in Self-regulated E-learning Environment, IEICE Information and Communication Technology Forum, IEICE Proceedings Series. 2014. doi:10.34385/proc.19.COMM1-2
38. Donthu N, Gustafsson A. Effects of COVID-19 on business and research. Journal of business research. 2020. 117, 284–289. doi:10.1016/j.jbusres.2020.06.008

39. Tuite AR, Zhu L, Fisman DN, Salomon JA. Alternative dose allocation strategies to increase benefits from constrained COVID-19 vaccine supply. Annals of internal medicine. 2020. PMID:33395334
40. Shrivastava SR, Shrivastava PS. Minimizing the risk of international spread of coronavirus disease 2019 (COVID-19) outbreak by targeting travelers. Journal of Acute Disease. 2020. Mar 1;9(2):47. doi:10.4103/2221-6189.278654
41. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN,..& Munster VJ. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. The New England Journal of Medicine. 2020; 382(16), 1564-1567. doi: 10.1056/NEJMc2004973
42. Centers for Disease Control and Prevention. Coronavirus disease 2019 - clinical care guidance. 2020. https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html. [accessed Dec 2, 2020].
43. National Institutes of Health. Covid-19 treatment guidelines. 2020. https://www.covid19treatmentguidelines.nih.gov/introduction/external. 2020. [accessed 12/1/20].
44. Soo RJJ, Chiew CJ, Ma S, Pung R, Lee V. Decreased influenza incidence under COVID-19 control measures, Singapore. Emerging infectious diseases, 2020; 26(8): 1933–1935. PMID:32339092
45. Akpan IJ, Akpan AA. Multiple criteria analysis of the popularity and growth of research and practice of visual analytics, and a forecast of the future trajectory. International Transactions in Operational Research. 2021. 28(5);1-24. doi:10.1111/itor.12952.
46. Masters-Waage TC, Jha N, Reb J. COVID-19, Coronavirus, Wuhan Virus, or China Virus? Understanding How to "Do No Harm" When Naming an Infectious Disease. Frontiers in Psychology, 2020. 1-11. doi:10.3389/fpsyg.2020.561270.
47. Jo S, Kim H, Kim S, Shin DH, Kim MS. Characteristics of flavonoids as potent MERS‐CoV 3C‐like protease inhibitors. Chemical biology & drug design. 2019. Dec;94(6):2023-30. PMID: 31436895
48. Xie M, Chen Q. Insight into 2019 novel coronavirus—An updated interim review and lessons from SARS-CoV and MERS-CoV. International Journal of Infectious Diseases. 2020 May 1; 94:119-24. 10.1016/j.ijid.2020.03.071
49. Mahase E. (2020). Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate. BMJ. 2020; 368. DOI:10.1136/bmj.m641.
50. World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that Causes it. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it. 2020. [accessed: Dec. 1, 2020].
51. World Health Organization. Coronavirus disease (COVID-2019) situation reports. 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019. [accessed: Dec 1, 2020].
52. Chary MA, Overbeek DL, Papadimoulis A, Sheroff A, Burns MM. Geospatial correlation between COVID-19 health misinformation and poisoning with household cleaners in the Greater Boston Area. Clinical toxicology. 2020 Sep 2:1-6. DOI:10.1080/15563650.2020.1811297.
53. Matias T, Dominski FH, & Marks DF. Human needs in COVID-19 isolation. Journal of Health Psychology, 2020; 25(7) 871–882. doi:10.1177/1359105320925149
54. Leung CC, Lam TH, Cheng KK. Mass masking in the COVID-19 epidemic: people need guidance. Lancet, 2020; 395(10228), 945. doi:10.1016/S0140-6736(20)30520-1
55. Roscoe RD, Grebitus C, O'Brian J, Johnson AC, Kula I. Online information search and decision making: Effects of web search stance. Computers in Human Behavior. 2016 Mar 1;56:103-18. doi:10.1016/j.chb.2015.11.028
56. Hölscher C, Strube G. Web search behavior of Internet experts and newbies. Computer networks. 2000; 33(1-6), 337-346. doi:10.1016/S1389-1286(00)00031-1
57. Krehbiel TC. SAS. In The Handbook of Software for Engineers and Scientists 2018. May 4;1047-1065. CRC Press. ISBN:9780203710012
58. Akpan IJ, Udoh EAP, Adebisi B. Small business awareness and adoption of state-of-the-art technologies in emerging and developing markets, and lessons from the COVID-19 pandemic. Journal of Small Business & Entrepreneurship, 2020; 1-18. doi:10.1080/08276331.2020.1820185
59. Lebow JL. The challenges of COVID-19 for divorcing and post-divorce families. Family process. 2020.

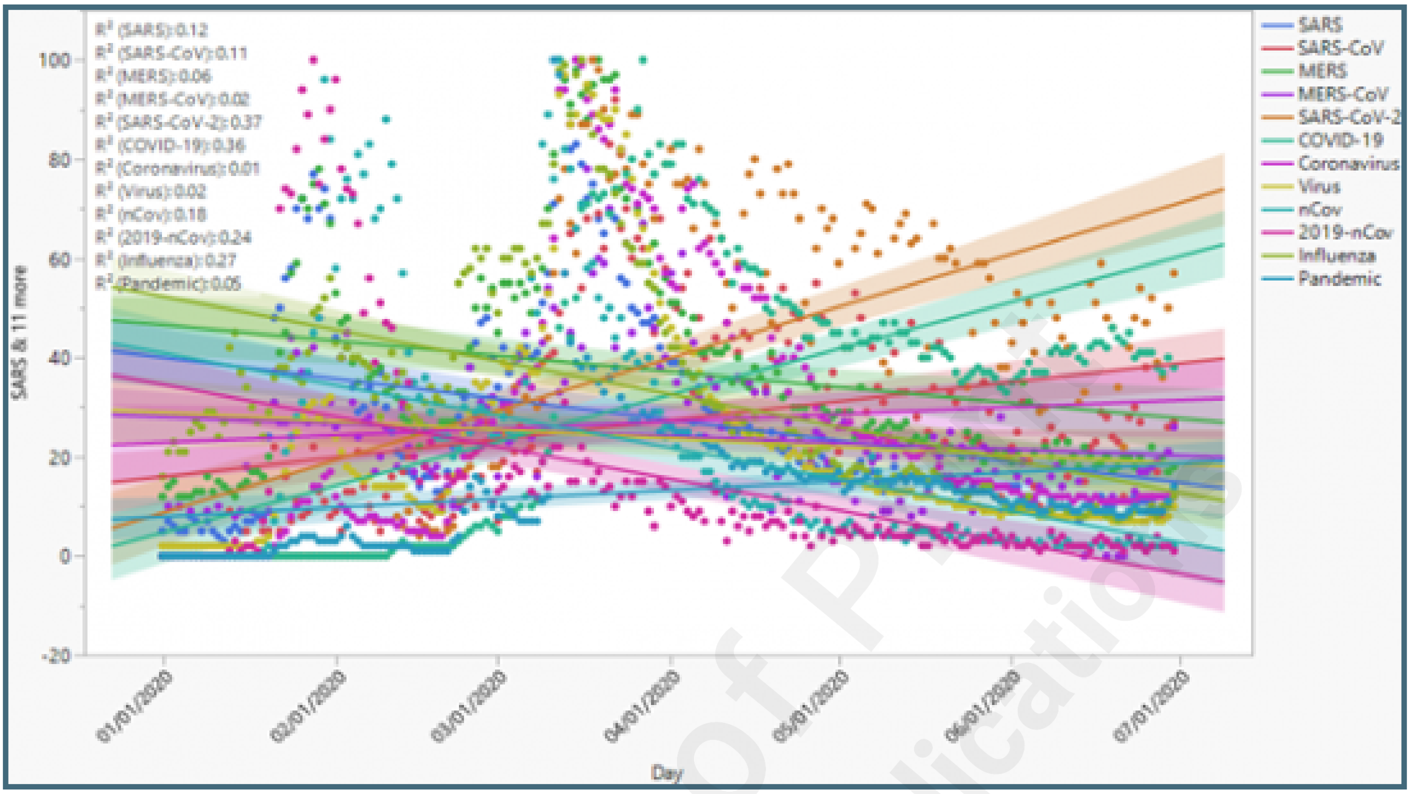
59(3):967-73. doi:10.1111/famp.12574

60. Akpan IJ, Ezeume IC. The Challenges Faced by Parents and Children from Divorce. Challenge. 2020 Nov 1; 63(6):365-77. doi:10.1080/05775132.2020.1842025
61. De Giorgio, A. (2020). COVID-19 is not just a flu. Learn from Italy and act now. Travel medicine and infectious disease. 2020 May-June; 35. doi:10.1016/j.tmaid.2020.101655
62. Brodeur A, Clark AE, Fleche S, Powdthavee N. COVID-19, lockdowns and well-being: Evidence from Google Trends. Journal of public economics. 2021 Jan 1;193. doi:10.1016/j.jpubeco.2020.104346
63. Sanchez-Caballero S, Selles MA, Peydro MA, Perez-Bernabeu E. An efficient COVID-19 prediction model validated with the cases of China, Italy, and Spain: Total or partial lockdowns? Journal of clinical medicine. 2020 May;9(5). doi.org/10.3390/jcm9051547
64. Akpan IJ, Shanker M, Razavi R. Improving the success of simulation projects using 3D visualization and virtual reality. Journal of the Operational Research Society. 2020. Dec 1. 71(12):1900-26. doi:10.1080/01605682.2019.1641649
65. Pickles K, Cvejic E, Nickel B, Copp T, Bonner C, Leask J, Ayre J, Batcup C, Cornell S, Dakin T, Dodd RH. COVID-19 Misinformation Trends in Australia: Prospective Longitudinal National Survey. J Med Internet Res. 2021 Jan 7;23(1). doi:10.2196/23805
66. Akpan IJ, Aguolu OG, Ezeume IC. Overcoming the challenge of communicating the concept and science of SARS-CoV-2 and COVID-19 to non-experts. Challenge, 2021; 64(2), 117-131. DOI:10.1080/05775132.2021.1912984.
67. Lima-Pereira P, Bermúdez-Tamayo C, Jasienska G. Use of the Internet as a source of health information amongst participants of antenatal classes. Journal of clinical nursing, 2012. 21(3-4), 322-330. doi:10.1111/j.1365-2702.2011.03910.x
68. Cervellin G, Comelli I, Lippi G. Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. Journal of epidemiology and global health, 2017. 7(3), 185. doi:10.1016/j.jegh.2017.06.001
69. Mavragani, A., & Ochoa, G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill. 2019 May 29; 5 (2): e13439. doi: 10.2196/13439.
70. Mavragani A, Ochoa G. The Internet and the Anti-Vaccine Movement: Tracking the 2017 EU Measles Outbreak. BDCC. 2018 Jan 16;2(1):2. doi:10.3390/bdcc2010002
71. Jun S, Yoo HS, Choi S. Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. Technological Forecasting and Social Change. 2018 May; 130;69-87. doi:10.1016/j.techfore.2017.11.009
72. Ko JY, Danielson ML, Town M, Derado G, Greenlund KJ, Kirley PD, et al. Risk Factors for Coronavirus Disease 2019 (COVID-19)–Associated Hospitalization: COVID-19–Associated Hospitalization Surveillance Network and Behavioral Risk Factor Surveillance System. Clin Infect Dis. 2021 Jun 19;72(11):e695–703. doi:10.1093/cid/ciaa1419
73. Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GHL, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ Open. 2021 Jan 1; 11(1). doi:10.1136/bmjopen-2020-044640
74. Akpan IJ, Akpan AA. The impact of internet use on students' learning outcomes in higher education in developing countries. International Journal of Education Research. 2017. 12(1);45. https://link.gale.com/apps/doc/A517626094/AONE?u=anon~8c2e4847&sid=bookmark-AONE&xid=c99ca4ab. [accessed January 2, 2020].
75. Mavragani A, Tsagarakis KP. Predicting referendum results in the Big Data Era. J Big Data 2019 Jan 14;6(3). doi:10.1186/s40537-018-0166-z

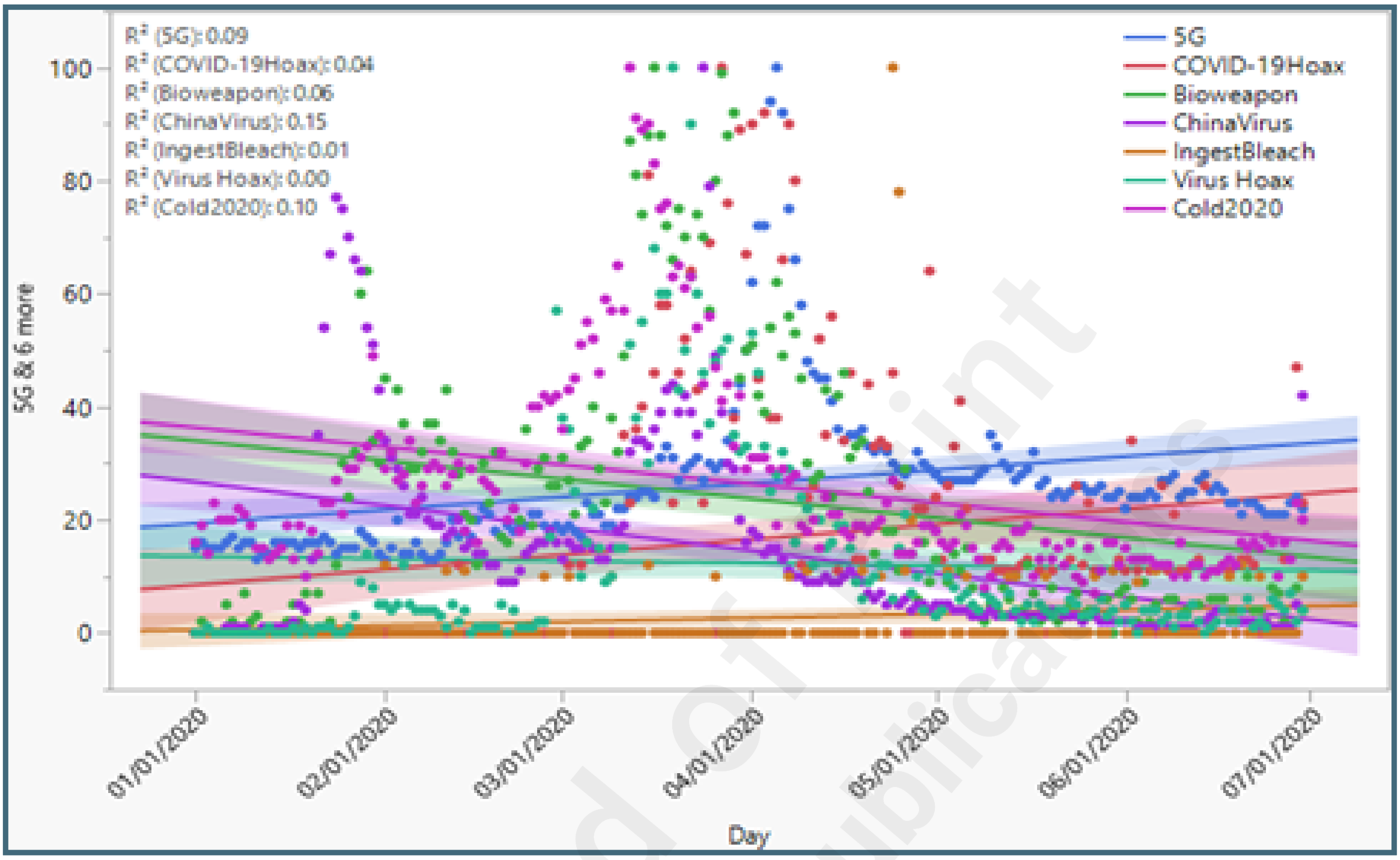
Supplementary Files

Figures

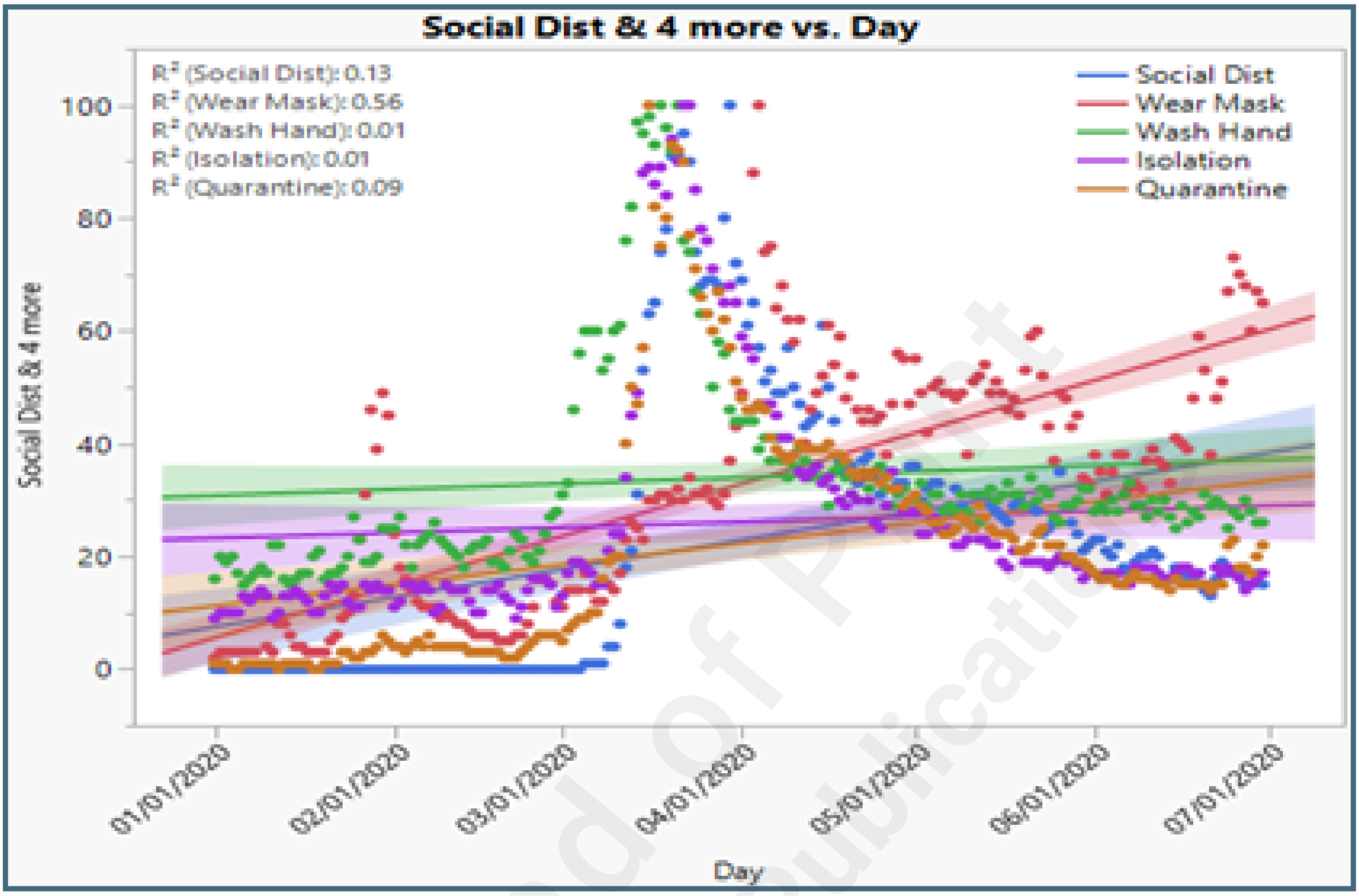
The keywords used by people to learn about SARS-CoV-2 and COVID-19 through we search.



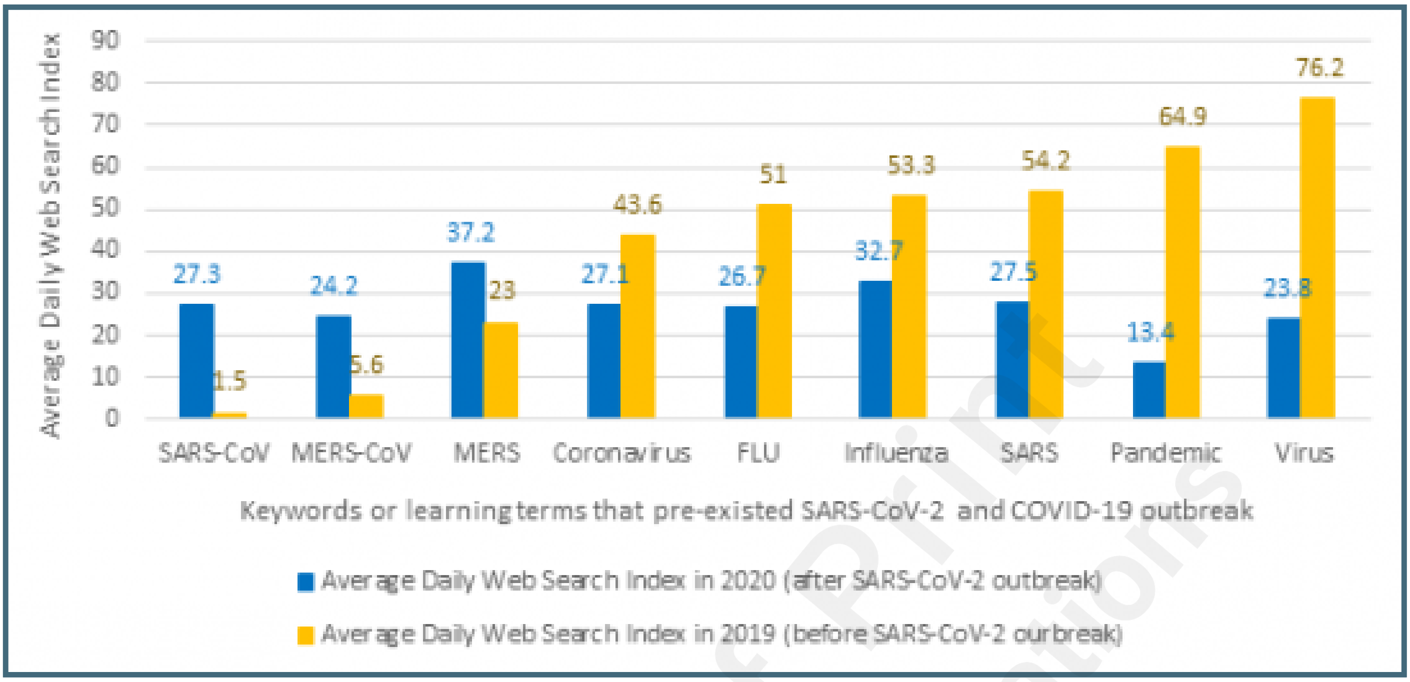
Worldwide search index showing the learning terms that represent misinformation and conspiracy theories about COVID-19.



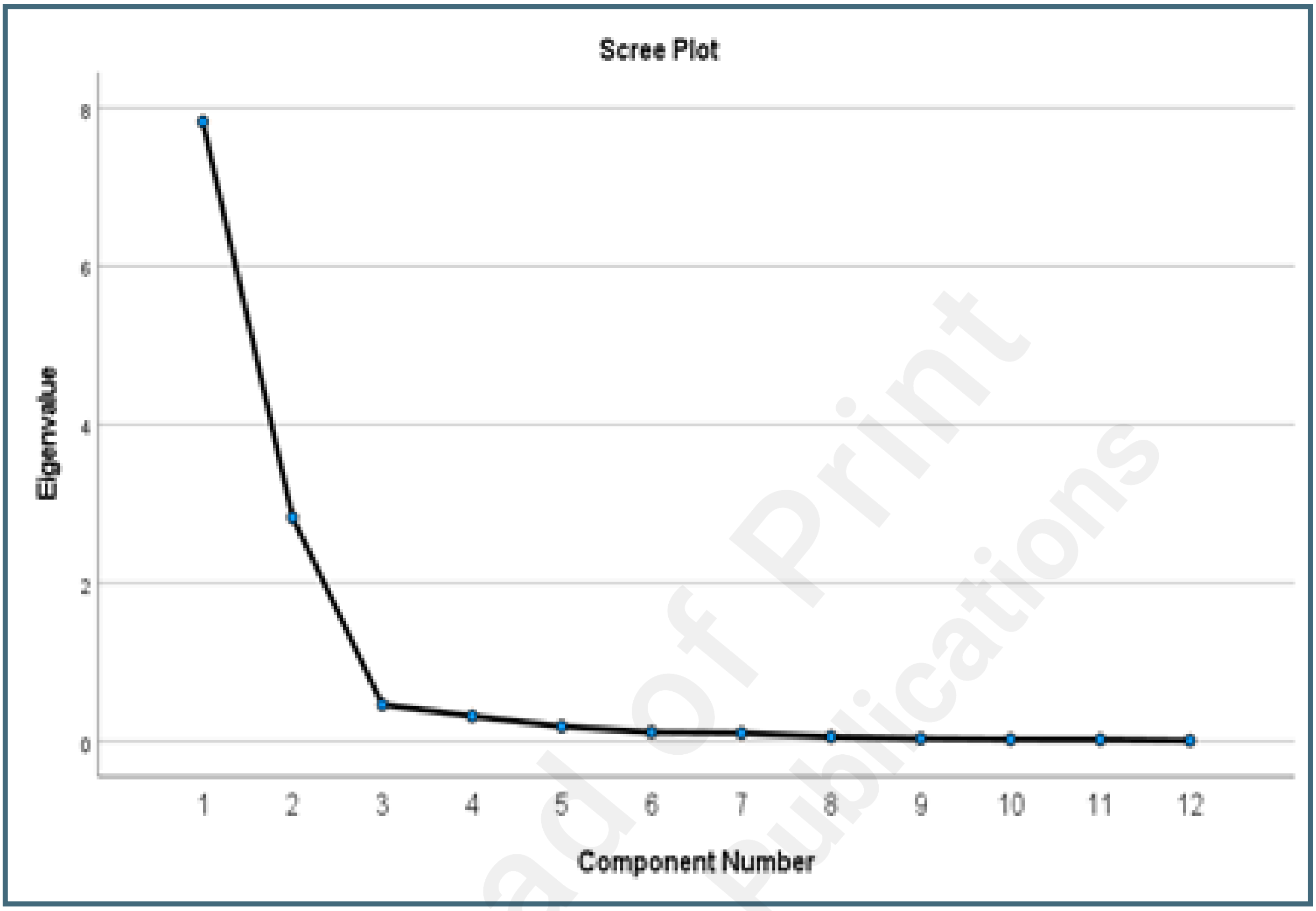
The trend analysis showing the web search index in learning about public health safety measures.



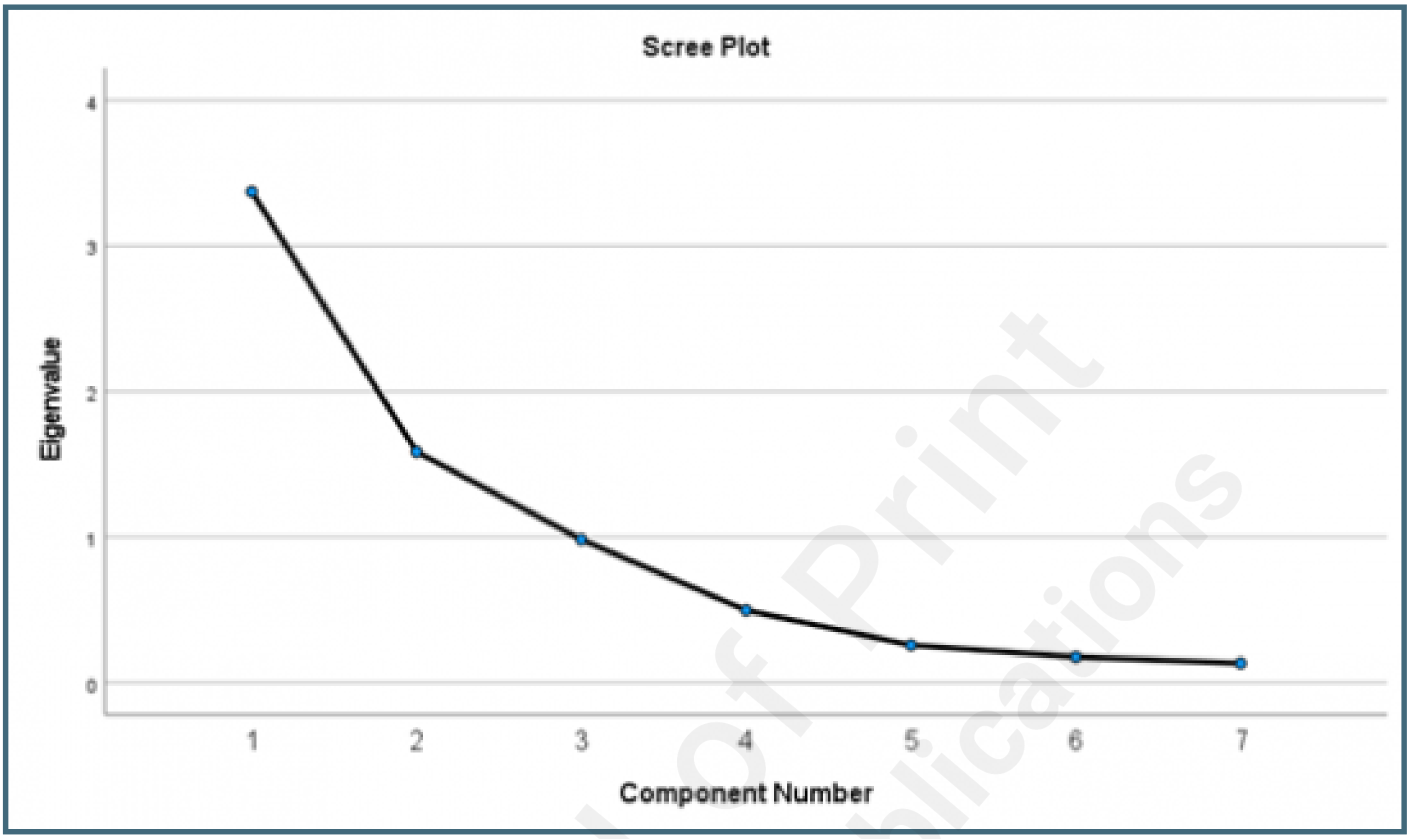
The average search index for pre-existed keywords before and during the COVID-19 pandemic.



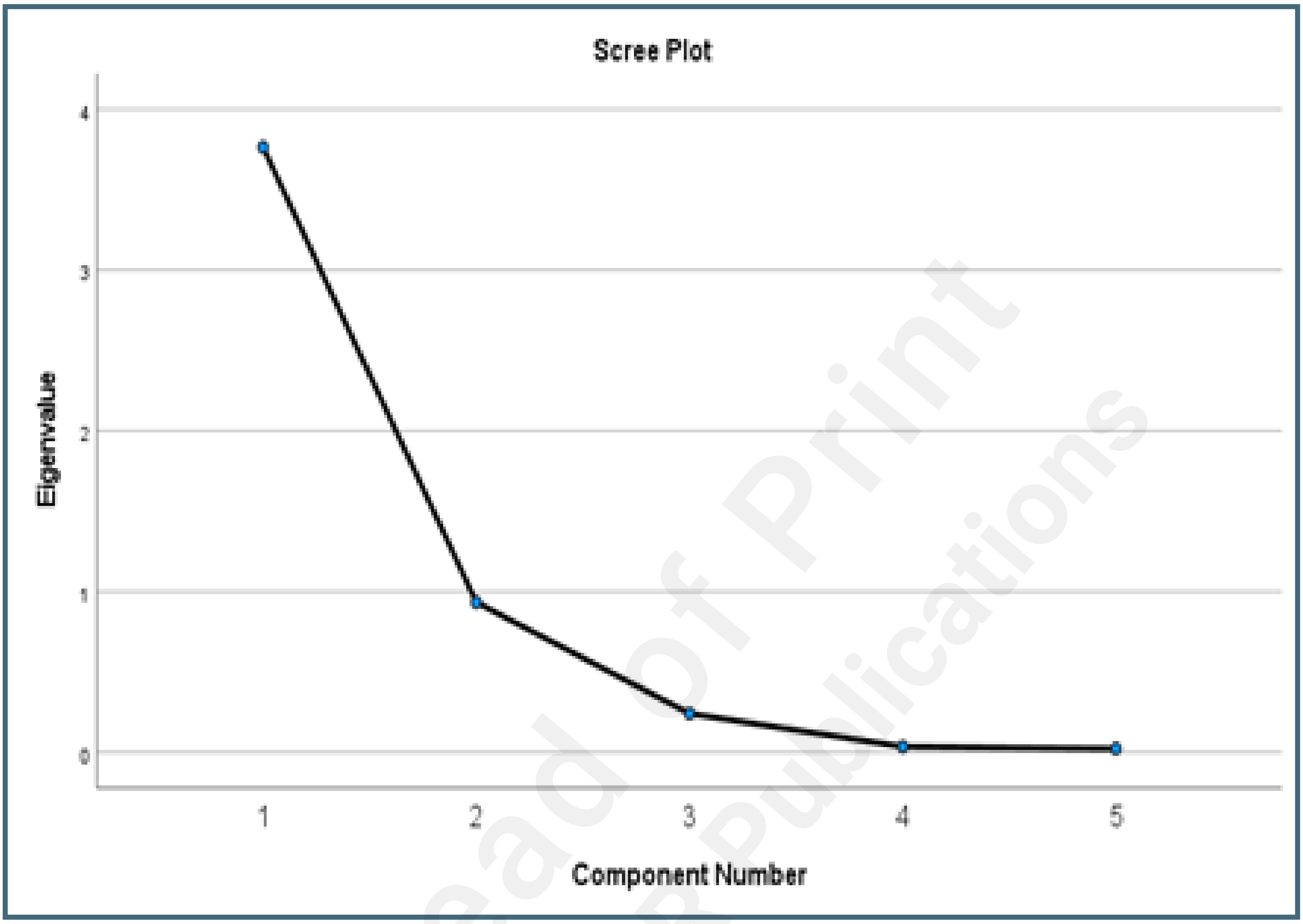
Scree plot of COVID-19 learning terms; two principal components extract 88.78% of total variation.



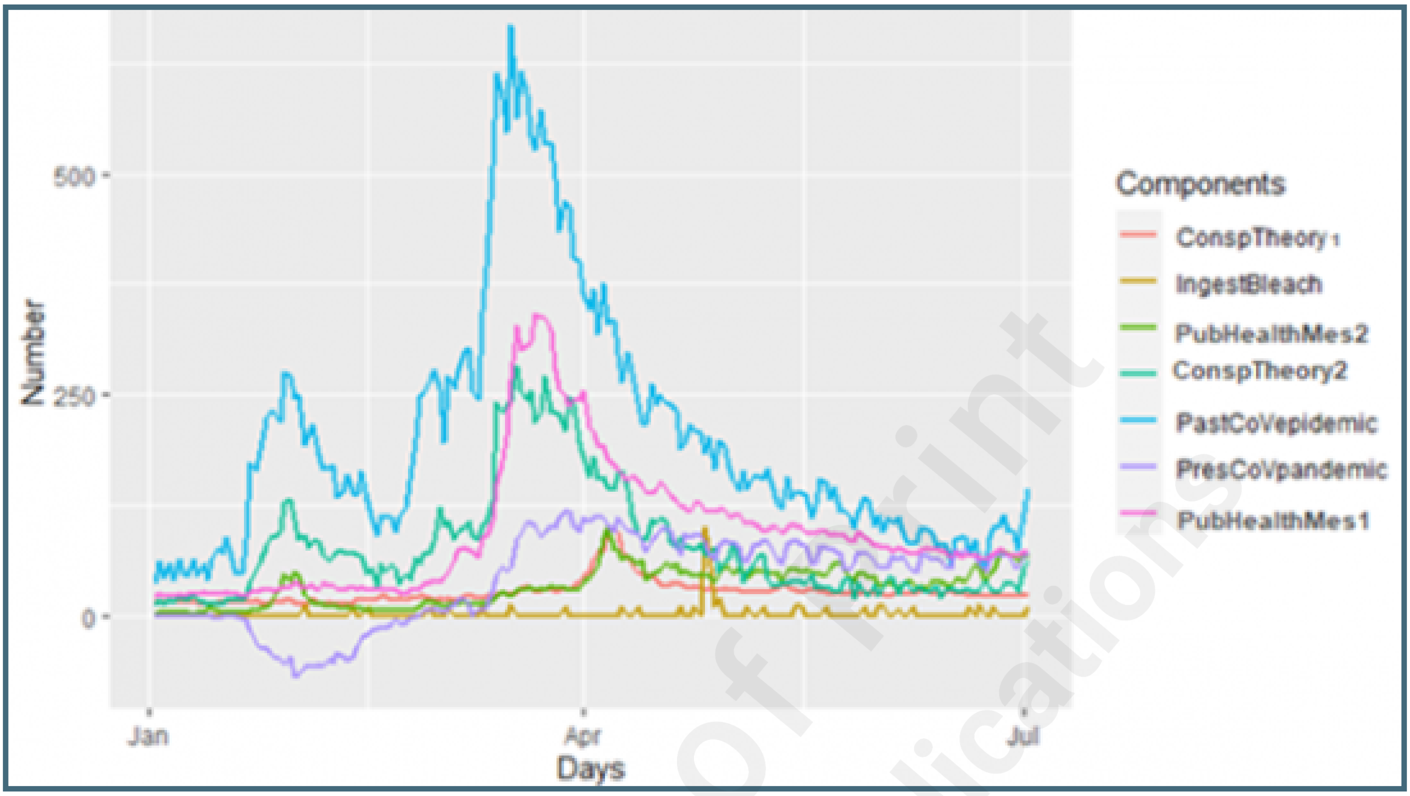
Scree plot of misinformation and conspiracy theory terms.



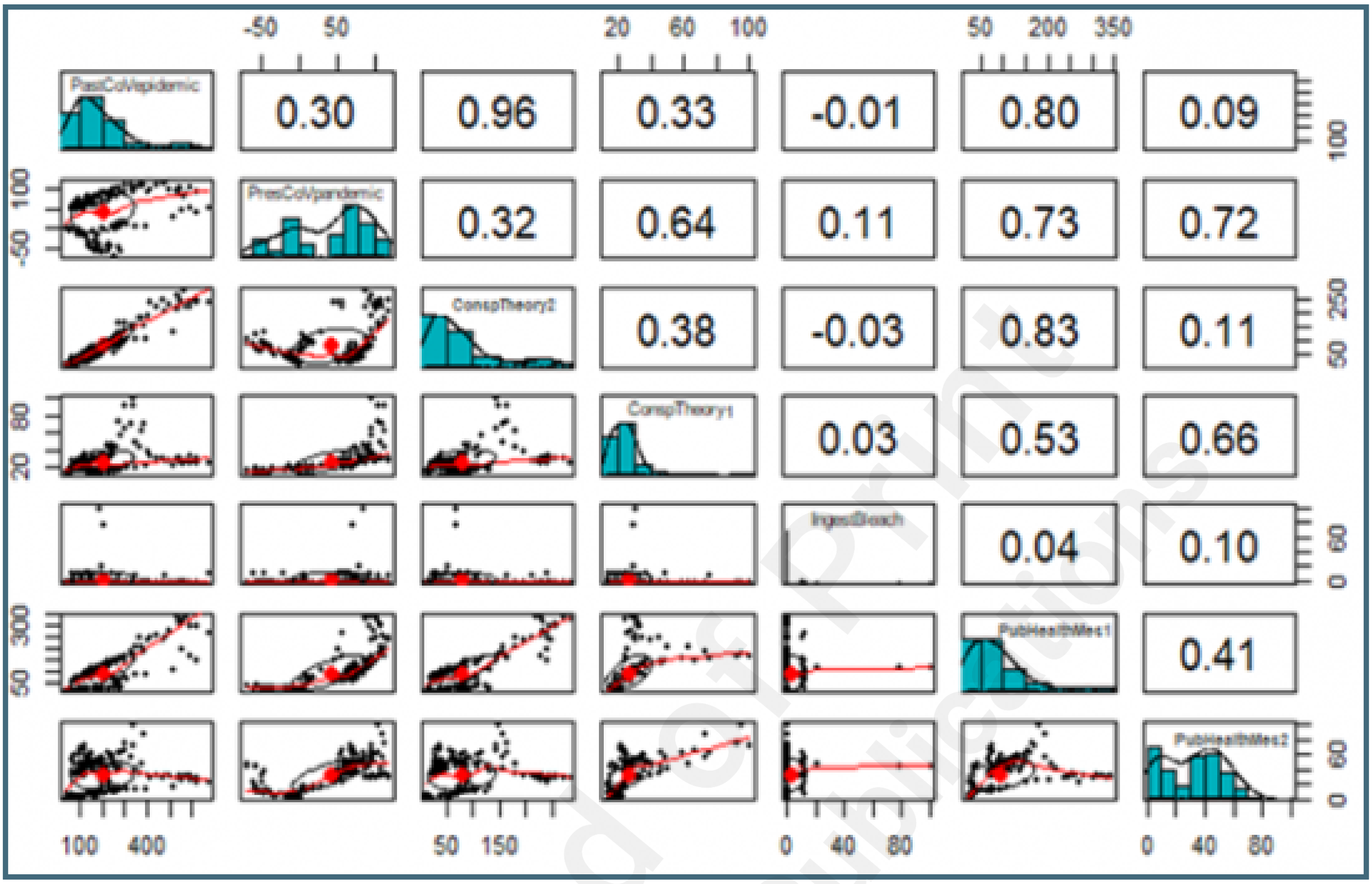
Scree plot of public health measures against COVID-19; two components extract 93.99% of the variability in the search index.



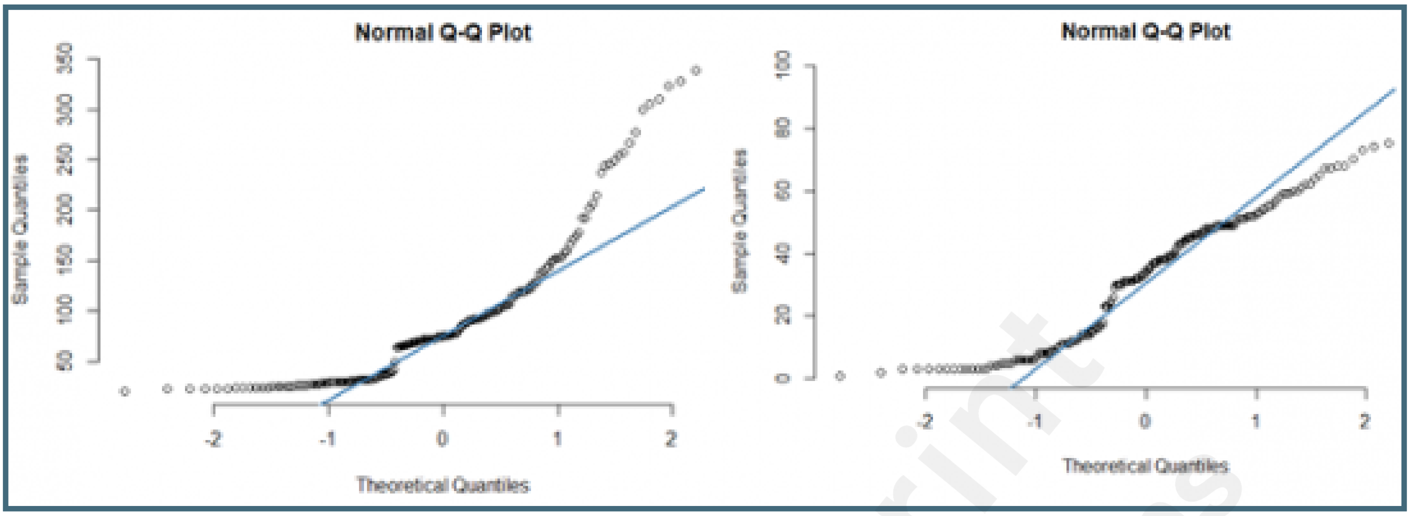
Daily search index of the principal components.



Panel Pair Plot of linear correlation.



Q-Q plot of PubHealthMes1 and PubHealthMes2.



Multimedia Appendixes

Table 2: Summary statistics of the normalized daily global Google Trend scores for different keywords used in this study. Data correspond to the time window between 1st Jan 2020 to 30th June 2020 (n=182).
URL: http://asset.jmir.pub/assets/2a8fbe92edd289a4a9d43fcfcf5b46ab.docx

CONSORT (or other) checklists

FOR EDITORS: Revised Manuscript with tracked changes and Authors' Responses to Reviewers' Feedback and Editorial Comments (please avoid duplication).
URL: http://asset.jmir.pub/assets/8e2f9ee1a8793608c181d2496f70915e.pdf