Research Articles

Using Google Trends to assess COVID-19 related web search behaviour in Arabian Peninsula

Aymane Ajbar, Mourad Boumaaza, Abdelhamid M Ajbar

1 Royal College of Surgeons in Ireland – Medical University of Bahrain, Bahrain, 2 Department of Chemical Engineering, College of Engineering, King Saud University, Riyadh, Saudi Arabia

Keywords: covid-19, arabian peninsula, google trends

Background

The relatively small, populated countries (Bahrain, Kuwait, Oman, Qatar, and the United Arab Emirates) of the Arabian Peninsula have recorded high per-capita COVID-19 infection rates. These countries also have a high internet penetration rate with a large Google share. Google Trends is a freely accessible online tool that allows the user to analyze trends in public Google searches during a given period and across a specified region. This paper investigates the use of this tool to study public awareness of COVID-19 in the selected countries.

Methods

The information sought by the public in the selected countries was grouped into three categories: (i) Coronavirus search item, (ii) COVID-19 symptoms, and (iii) protection measures. Data from Google Trends were analyzed using Joinpoint software. This approach enabled searches with a meaningful trend to be delineated and allowed them to be ranked by order of importance.

Results

Overall, public interest in COVID-19 was found to be associated with seven symptoms (fever, cough, sore throat, headache, loss of smell, shortness of breath, and chest pain), and four preventive measures (social distancing, quarantine, hand sanitizers, and masks). However, Google trends found some disparities between the studied countries.

Conclusions

Despite its limitations, Google Trends can be useful in uncovering the major areas of public interest in COVID-19 for countries with high internet penetrance rates. Data from Google Trends could be used as part of a health communication strategy to identify information needs and establish appropriate communication strategies to tackle the pandemic.
Using Google Trends to assess COVID-19 related web search behaviour in Arabian Peninsula

Table 1. Some characteristics of the studied countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Population (2020)</th>
<th>Number of COVID-19 per 1,000 people (Dec 30, 2020)</th>
<th>Internet penetration rate (2020)</th>
<th>Google share (July 2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bahrain</td>
<td>1,701,575</td>
<td>55.5</td>
<td>94.9 %</td>
<td>97.5%</td>
</tr>
<tr>
<td>Kuwait</td>
<td>4,270,571</td>
<td>36.2</td>
<td>99.1 %</td>
<td>97.4%</td>
</tr>
<tr>
<td>Oman</td>
<td>5,106,626</td>
<td>25.3</td>
<td>78.5 %</td>
<td>96.8%</td>
</tr>
<tr>
<td>Qatar</td>
<td>2,881,053</td>
<td>52.1</td>
<td>99.6 %</td>
<td>97.6%</td>
</tr>
<tr>
<td>UAE</td>
<td>9,890,402</td>
<td>24.0</td>
<td>96.4 %</td>
<td>96.6%</td>
</tr>
</tbody>
</table>

UAE – United Arab Emirates. Source: 1,12,15

Table 1 shows the characteristics of the studied countries in the Arabian Peninsula. The table includes population, number of COVID-19 cases per 1,000 people, internet penetration rate, and Google share for July 2020.

METHODS

SETTINGS

The Arabian Peninsula consists of the countries of Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, UAE, and Yemen. All these countries are similar in terms of culture. Yemen has a much smaller GDP per capita, lower internet penetration rate, and the country is engulfed in a civil war. Therefore, it was excluded from this study. Saudi Arabia, with a population of approximately 34 million,15 is by far the largest country in the region, exceeding the other countries by many orders of magnitude (Table 1). To maintain an unbiased comparison, we also have excluded Saudi Arabia from our analysis. Therefore, we have limited our analysis to the five countries of Bahrain, Kuwait, Oman, Qatar, and UAE. All are relatively small countries with respect to population size. Regarding COVID-19 cases, these countries have reported high ratios (Table 1), with Bahrain announcing, as of December 30, 2020, a total of 55.5 cases per thousand people, followed by Qatar (52.1), Kuwait (36.2), UAE (24.0), and Oman (25.5). Moreover, with the exception of Oman, all other countries have very high internet penetration rates of at least 94.9% (Table 1). The Google internet share in these countries is also similar and is no less than 96% (Table 1). Given these facts, GT data for these countries should typically be helpful in revealing major trends of public interest in the disease, its symptoms, and the inclination of the population to take protective measures.

DATA COLLECTION

GT collects data from all Google searches, with the popularity of a keyword determined by the relative search volume (RSV). When a search term is selected, a "search category" is chosen that can incorporate all variations of the search term as well as different languages. This approach is important because Arabic is the dominant language in these countries, along with other languages spoken by a large number of expatriates.

Because internet searches associated with COVID-19 can cover a wide range of topics, we have limited our investigation to three search areas: (a) searches for "Coronavirus", (b) searches for the disease symptoms, and (c) searches for protective measures to control the spread of the disease.

The COVID-19 symptoms were taken from the World Health Organization (WHO) website.16 The list contains the most common symptoms (fever, cough, and tiredness), less common symptoms (aches and pains, sore throat, diarrhea, conjunctivitis, headache, and loss of smell), and serious symptoms (shortness of breath and chest pain). The list of preventive measures consists of social distancing, quarantine, hand sanitizers, and masks.

DATA ANALYSIS

For each search keyword, the RSV data from GT were extracted for the period extending from January 1, 2020 to August 8, 2020. The reason we restricted the analysis to this period is that the interest in Coronavirus faded after this period, as will be discussed later. Data were subjected to statistical analysis using the Joinpoint regression program18 to look for the existence of statistically significant trends. Such trends are referred to as "joinpoints"—a zero joinpoint would indicate no change in trend. The joinpoint methodology is better suited here to uncover statistically significant changes in a trend, which other methods (e.g., autoregressive integrated moving average analysis) may miss.19

Joinpoint analysis has been used successfully with and without GT data in a number of health studies.8,20 The joinpoint analysis allowed us to exclude non-meaningful search items (i.e., those that did not display any significant changes in trend). The average RSV values, on the other hand, allowed to rank the search items by order of importance.

RESULTS

CORONAVIRUS SEARCH INTEREST

The RSV results for the "Coronavirus" search in the five countries for the selected time period are shown in Figure 1. The RSV for each country is calculated by dividing the sampled search data points by the total number of searches done in that specific country and during the given time period, after which the resulting values are scaled from 0 to
100. All countries (Figure 1) show more or less a similar trend that consists of an increase in RSV, reaching the peak of 100 and then fading out. Table 2 summarizes the quantitative results for the "Coronavirus" search. In terms of relative search volume, it can be seen that the average value of RSV for Qatar is 30.1 and is larger than Bahrain (26.7), UAE (26.5), Oman (22.8), and Kuwait (21.6). The data were then fed into the Joinpoint regression program for analysis. The program began with the minimum number of joinpoints (0, which corresponds to the simple linear regression model) and checked whether more joinpoints must be added to the model. The results of the data fitting are shown in the joinpoint graph (Figure 2) for UAE (for the sake of space limitations, the figures for other countries are not included because they have a similar trend). The y-axis of the figure shows the RSV data (labeled by the program as "proportion"), whereas the x-axis represents the date order, starting from January 1, 2020. The results of the figure (legend on the right) indicate that the program managed to find three joinpoints and Figure 2, therefore, divided the period under study into four time segments that are statistically significant.

The existence of three joinpoints for each country indicates a real change in trend in the RSV data. Table 2 shows the locations (in days) of the joinpoints for each country. Some features of these points are analyzed in a later section.

DISEASE SYMPTOMS

Table 3 shows the RSV results for the disease symptoms. It can be seen that for all these countries, no joinpoints (and, therefore, no change in trend) were found for fatigue, diarrhea, conjunctivitis, and myalgia symptoms. Apparently, these four symptoms did not receive the attention of Google users in any of the studied countries. Fever and cough, on the other hand, showed a meaningful trend (2 joinpoints) in all the studied countries, except Oman. Sore throat was meaningful only in Kuwait, Qatar, and UAE. Searches for the headache symptom were sought only in UAE, loss of smell was meaningful only in Kuwait and UAE (2 joinpoints), shortness of breath was searched for in Kuwait, Qatar, and UAE, and finally, searches for chest pain were meaningful only in UAE. Figure 3 shows an example of RSV plots over the studied period for fever in all countries except Oman, where searching for fever was not statistically significant.

Table 2. Joinpoint results of "coronavirus" search

<table>
<thead>
<tr>
<th>Country</th>
<th>Bahrain</th>
<th>Kuwait</th>
<th>Oman</th>
<th>Qatar</th>
<th>UAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average RSV</td>
<td>26.0</td>
<td>21.6</td>
<td>22.8</td>
<td>30.1</td>
<td>26.5</td>
</tr>
<tr>
<td>Location of the three joinpoints (days)</td>
<td>18/2/2020</td>
<td>18/2/2020</td>
<td>9/3/2020</td>
<td>22/2/2020</td>
<td>27/2/2020</td>
</tr>
</tbody>
</table>

RSV = relative search volume, UAE = United Arab Emirates.

PROTECTIVE MEASURES

Table 4 shows that apart from Bahrain, all other countries experienced a meaningful search for all the four selected prevention measures (social distancing, quarantine, hand sanitizers, and masks). For Bahrain, the users were interested only in the quarantine (average RSV of 11.1). For Kuwait, the order of importance was as follows: quarantine (14.5), masks (12.8), hand sanitizers (9.0), and social distancing (7.6). For Oman, the order of importance was: masks (28.0), quarantine (17.2), social distancing (7.4), and hand sanitizers (7.2). For Qatar, the order of importance was similar to Oman: masks (21.5), quarantine (19.6), social distancing (11.8), and hand sanitizers (7.1). Finally, for UAE, the order of importance was: quarantine (23.3), social distancing (17.5), hand sanitizers (14.3), and masks (11.8).

DISCUSSION

The significant media coverage surrounding the COVID-19 pandemic has certainly generated substantial web search activities in Google. This phenomenon explains why Google users in the studied countries began searching for Coronavirus around the time their countries reported their first cases. In this regard, the first COVID-19 cases in the five countries were reported around the same time: Bahrain (February 21), Kuwait (February 24), Oman (February 24), Qatar (February 29), and UAE (February 29). The other hand, shows that the first joinpoint was observed respectively on February 18 for Bahrain and Kuwait, March 9 for Oman, February 22 for Qatar, and February 27 for UAE. These dates are close to the days of the first reported cases, after which there was a sharp increase in search volume (Figure 1). Hence, there was a strong, positive relationship between public interest and the growing disease burden faced by these countries. The second joinpoint (Figure 2 and Table 2) corresponds, on the other hand, to the peak of search volumes. It can be seen (Table 2) that search interest in all the countries reached a peak around the same time: April 4 to April 9. After the peak, a fall in trend of the extracted RSV was seen (e.g., Figure 1) despite the increase with time of COVID-19 cases in these countries. This finding is most likely due to the public becoming familiar with the pandemic and, therefore, searches for it were less frequent.

Public interest in COVID-19 symptoms in the selected countries was overall found to be associated with the following seven symptoms: fever, cough, sore throat,
### Table 3. Joinpoint analysis for disease symptoms

<table>
<thead>
<tr>
<th>Country</th>
<th>Bahrain</th>
<th>Kuwait</th>
<th>Oman</th>
<th>Qatar</th>
<th>UAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
</tr>
<tr>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
</tr>
<tr>
<td>Fever</td>
<td>16.0</td>
<td>35.2</td>
<td>-</td>
<td>27.5</td>
<td>30.8</td>
</tr>
<tr>
<td>Cough</td>
<td>22.5</td>
<td>40.2</td>
<td>-</td>
<td>30.1</td>
<td>34.8</td>
</tr>
<tr>
<td>Fatigue</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sore throat</td>
<td>-</td>
<td>26.7</td>
<td>-</td>
<td>30.3</td>
<td>29.6</td>
</tr>
<tr>
<td>Headache</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>55.7</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Myalgia</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Loss of smell</td>
<td>-</td>
<td>7.1</td>
<td>-</td>
<td>-</td>
<td>8.7</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>-</td>
<td>33.0</td>
<td>-</td>
<td>18.3</td>
<td>36.1</td>
</tr>
<tr>
<td>Chest pain</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>44.1</td>
</tr>
</tbody>
</table>

RSV – relative search volume, UAE – United Arab Emirates. Average RSV are calculated only for nonzero joinpoints.

### Table 4. Joinpoint analysis for protection measures

<table>
<thead>
<tr>
<th>Country</th>
<th>Bahrain</th>
<th>Kuwait</th>
<th>Oman</th>
<th>Qatar</th>
<th>UAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
<td>Average RSV</td>
</tr>
<tr>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
<td>Numbe of joinpoints</td>
</tr>
<tr>
<td>Social distancing</td>
<td>-</td>
<td>7.6</td>
<td>7.4</td>
<td>11.8</td>
<td>17.5</td>
</tr>
<tr>
<td>Quarantine</td>
<td>11.1</td>
<td>14.5</td>
<td>17.2</td>
<td>19.6</td>
<td>23.3</td>
</tr>
<tr>
<td>Hand sanitizer</td>
<td>-</td>
<td>9.0</td>
<td>7.2</td>
<td>7.1</td>
<td>14.3</td>
</tr>
<tr>
<td>Mask</td>
<td>-</td>
<td>12.8</td>
<td>28.0</td>
<td>21.3</td>
<td>11.8</td>
</tr>
</tbody>
</table>

RSV – relative search volume, UAE – United Arab Emirates. Average RSV are calculated only for nonzero joinpoints.
headache, shortness of breath, chest pain, and loss of smell. However, there are disparities between the studied countries. Bahrain reported interest only in two symptoms, Kuwait five, Qatar four, whereas UAE reported an interest in all the symptoms. Oman, on the other hand, did not report interest in any of the symptoms. There are three plausible explanations for these disparities: One has to do with the low internet penetrance rate in Oman (78%) compared to the other countries (>95%). The other reason is that the awareness of the disease was probably more wide-
spread in social media in some countries; thus, interested people could have sought online health information using social media, such as Facebook or Twitter. In this regard, it was shown that social media has a major influence on health-seeking behavior in some countries of the Arabian Peninsula. These disparities may also be due to poor data collection by GT in Bahrain and Oman. Some extracted GT data for these two countries showed a high variability in RSV, producing alternating zero values (suggesting that no searches had been undertaken on that day). It could be that a significant proportion of users were executing useful searches during those dates; however, GT labeled low volume data relative to the peak as zero.

Moreover, in terms of the importance of the volume of searches, it can be seen from Table 3 that for Bahrain, searches for cough were more important (average RSV of 22.5) than fever (16.0). For Kuwait, the order of importance was the following: cough (average RSV of 40.2), fever (35.2), shortness of breath (35.0), sore throat (26.7), and loss of smell (7.1). For Qatar, the order of importance was: sore throat (50.3), cough (50.1), fever (27.5), and shortness of breath (18.3). Finally, for UAE, the order of importance was: headache (55.7), chest pain (44.1), shortness of breath (36.1), cough (34.8), fever (30.8), sore throat (29.6), and loss of smell (8.7). One reason why some symptoms were searched for more than others in any country could be due to the fact that they reflect the actual symptoms of COVID-19 in that specific country. However, because COVID-19 epidemiological data are not being released by the countries’ health authorities, it is not possible to conclude with certainty that the symptoms found through GT are the actual prevalent disease manifestations. However, the GT results are quite useful at this point and show at least the major areas of interest of Google users in COVID-19 symptoms in each country.

Figure 4 shows the variations of the RSV data for the loss of smell symptom for Kuwait and UAE. The search for this symptom was statistically significant (i.e., 2 joinpoints) only for these two countries, although fewer people searched for it as judged by the small average RSV values shown in Table 3. The first joinpoint for UAE was detected on April 14. It should be noted that the earliest date when loss of smell was added to the list of official symptoms of COVID-19 was on April 26 by the US CDC. Here, we see a situation where people in UAE searched on the internet for a COVID-19 symptom long before it was added officially to the list of symptoms, a result consistent with the conclusions reached in, who used GT to investigate searches on the loss of smell symptom associated with COVID-19 in eight countries.

With respect to prevention measures to control the spread of the disease, quarantine seemed to be searched for in all the studied countries. Moreover, the other protective measures (social distancing, masks, and hand sanitizers) were also searched for in all the countries except Bahrain. The most likely reason for this disparity is again poor data collection and management by GT because the extracted GT data for Bahrain showed alternating zero values. It may be possible that many users were executing useful searches, but GT ignored low volume data relative to the peak.

CONCLUSIONS

This study is the first of its kind to assess public awareness of COVID-19 in the five relatively small countries of the Arabian Peninsula. The results of this study are strengthened by the particularly high internet penetration rate in four of the five studied countries, the high market share of Google, and the inclusion of search keywords in different languages. From this point of view, Google Trends was useful in uncovering the major areas of public interest associated with the COVID-19 pandemic. In this regard, GT was found to be strongly used in connection with COVID-19 in UAE, Kuwait, and Qatar. However, GT data were found to have limited use in Bahrain and especially in Oman. This difference may be due to the low internet penetration rate in Oman, poor data collection where GT ignored small search volumes, and most likely the preference of people to use social media or other venues for their inquiries. The use of GT in this study is important for understanding online health information seeking behavior in these countries.
of Google Trends as a public health tool could be substantially improved if its methods for data collection are publicly known. Data mining of the public is a growing concern. However, the data collected by GT is publicly available, completely anonymous, and user activity is only monitored when using the Google Search Engine and not their internet activity afterward. Finally, our methods could be applied to other regions, which may show different results. This strategy could offer insight into the behavior of the SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) with respect to different ethnicities and climates, ultimately allowing improved quality of information being released to the public. This approach will also allow accurate monitoring of the disease outbreak and public response.

FUNDING

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group no (RG-1441-188)

AUTHORSHIP CONTRIBUTIONS

AA: Initial study design, collection of data, results analysis and manuscript draft.
MB: Data analysis and review of the manuscript
AMA: Data analysis and review of manuscript
All authors approved the final version of the manuscript.
AMA is the guarantor.

COMPETING INTERESTS

All authors have completed the Unified Competing Interest form available at http://www.icmje.org/conflicts-of-interest/ in line with the Journal of Global Health editorial policy and declare no conflicts of interest.

CORRESPONDENCE TO:

Abdelhamid Ajbar, Ph.D, College of engineering. King Saud University, Riyadh, Saudi Arabia; Email: aajbar@ksu.edu.sa

Submitted: January 14, 2021 GMT, Accepted: February 15, 2021 GMT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at http://creativecommons.org/licenses/by/4.0 and legal code at http://creativecommons.org/licenses/by/4.0/legalcode for more information.
REFERENCES


Using Google Trends to assess COVID-19 related web search behaviour in Arabian Peninsula