Analysis and potential use of Google Trends as a monitoring tool for risk communication during COVID-19 pandemic

Nindya Mahfuza¹, Rizma Adlia Syakurah², Resiana Citra³
¹²³Medical Faculty, Sriwijaya University, Indonesia

ABSTRACT

COVID-19 Pandemic has become a major problem in various infected countries, including Indonesia. The proper risk communication strategy during this outbreak was important to reduce the impact. Therefore, this research was intended to assess the potential use of Google Trends as a tool to monitor risk communication during COVID-19 pandemic in Indonesia. Search patterns were analyzed using the terminology used to identify COVID-19 in Indonesia, followed by information-finding keywords 'gejala (symptoms)', 'mencegah (preventing)', and 'obat (drug)' keywords compared to the number of newly confirmed COVID-19 cases in Indonesia using time-lagged correlation analysis from December 31th, 2019 to April 20th, 2020. Peaks within respective timelines were qualitatively described according to current COVID-19 related events. "Corona" was the terminology mostly used in Indonesia to identify COVID-19. There were five spikes observed from “corona” keyword timeline, which each spike was dependent on the media coverage and regulation by the Government. Validation using time-lagged correlation yields significant results between corona, corona symptoms, preventing corona and corona drugs compared to newly confirmed COVID-19 cases in Indonesia. Google Trends can potentially be used to maximize the improvement of risk communication and as a tool to monitor public restlessness during the COVID-19 pandemic in Indonesia by Government.

Keywords:
COVID-19, Google Trends, Indonesia, Risk communication

This is an open access article under the CC BY-SA license.

Corresponding Author:
Rizma Adlia Syakurah, Public Health Faculty, Sriwijaya University, Jl. Raya Palembang-Prabumulih KM. 32 Indralaya, Ogan Ilir, Sumatera Selatan, 30662, Indonesia. Email: rizma.syakurah@gmail.com

1. INTRODUCTION

Since WHO declared corona virus outbreak as a global health emergency on January 30th, 2020, the number of corona virus disease 2019 (COVID-19) cases in the world has been rising. Within 2 weeks, the case of COVID-19 outside China has significantly increased 13-fold in the number of infected countries, on March 11th, 2020; there were over 118,000 cases of COVID-19 in 114 countries and 4,291 deaths. Therefore, WHO categorizes COVID-19 as a pandemic [1]. This pandemic has become a major problem in many affected countries, including Indonesia. In Indonesia, the first COVID-19 case was reported on March 2nd, 2020, subsequently reaching 4,839 positive cases, 426 cases of recovery and 456 cases of death by April 14th, 2020 [2].

Risk communication is used to view the dissemination of information among the general public regarding health risks and incidents, such as a disease outbreak as well as instructions on how to change
behavior to reduce the impact [3,4]. One of the lessons learned from major health problems of the 21st century according to WHO such as SARS, MERS, Influenza (H1N1), and Ebola is that risk communication and community engagement (RCCE) is an integration of successful responses to emergency health issues [5]. Risk communication is important in this COVID-19 pandemic. Assessing public restlessness, public response to information and public behavior changes during a pandemic is important to observe, given that the availability and accessibility of technology means that there is an increased likelihood of the presence and transmission of incorrect information, the spread of rumors and misinformation [3]. Therefore, robust risk communication is needed, particularly in affected countries, to assess public restlessness, public response to information, and public behavior changes during this COVID-19 pandemic [6].

The use of the Internet as health service resources is increasing. Infodemiology and infовeillance are public health information methods used to analyze search habits on the Internet [7]. One tool that can be used to facilitate researchers in obtaining Internet search data is Google Trends, an online portal that can be freely accessed anywhere [8]. Therefore, researchers are interested in using Google Trends data for monitoring how the Indonesian community response in finding data of COVID-19, so it can be potentially used as an early warning system and monitoring tool for public reactions, then can be used to help the risk communication during COVID-19 pandemic in Indonesia.

2. RESEARCH METHOD

The keywords that define COVID-19 in Indonesia were searched using Google Trends (GT) (https://trends.google.com/trends). The search was performed from December 31th, 2019 to April 20th, 2020. The selected keywords were ‘COVID’, ‘coronavirus’, and ‘corona’. Then we analyzed every spike that occurred, after obtaining the terminology that defines COVID-19 in Indonesia, we added ‘gejala (symptoms)’, ‘mencegah (prevent)’, and ‘obat (drug)’ keywords to assess the curiosity of people in Indonesia whether based on diagnosis (gejalal symptoms), prevention (mencegah/ preventing) or therapy (obat/ drugs).

The highest concern in a keyword was stated in the number 100, and the lack of attention or data was expressed in 0. GT data then compared with the daily data of COVID-19 cases collected from the official website of the acceleration Task Force Handlers COVID-19 (Acceleration of Handling Task Force COVID-19/Gugus Tugas Percepatan Penanganan COVID-19) in Indonesia (www.COVID19.go.id). Time-lagged correlation with a three-day interval of GT data and the number of newly confirmed COVID-19 cases was analyzed using Pearson Correlation with significance ≤ 0.05 to assess whether the increase of GT data was correlated with the increase of COVID-19 cases, as has been applied in other studies [9].

3. RESULTS AND DISCUSSION

Indonesia GT relative search volume (RSV) for ‘Covid, Coronavirus and Corona’ from December 31th 2019 to April 20th 2020 can be observed in Figure 1. The results show that terminology used to define COVID-19 in Indonesia was ‘corona’ compared to worldwide that used ‘coronavirus’ as a commonly used terminology as shown in Figure 2. Based on Google Trends data, corona-related searches in Indonesia were extremely fluctuative. There were five peaks for corona in Indonesia as follow: January 27th, March 2nd, March 16th, March 22nd, and April 4th, 2020. The highest peak occurred on March 22nd, 2020, when the Government announced positive cases of COVID-19 rapidly reached 514 people.

Data regarding queries related to keywords search COVID, coronavirus, and corona are shown in Table 1. Based on existing data, the searches in Indonesia were generally about the characteristics of coronavirus and coronavirus symptoms. Table 2 shows the distribution of cities with the most searched activity about COVID, coronavirus and corona. It was found Bekasi city, Tangerang and Depok as the most popular city in the search activity for COVID-19 in Indonesia. Figure 3 shows the comparative Indonesia GT RSV for ‘gejala (symptom) corona, mencegah (preventing) corona, and obat (drug) corona’ with newly confirmed COVID-19 cases. The result found that the three keywords have a similar pattern. There are 4 peaks on January 27th, 2020, March 2nd, 2020, March 16th, 2020, and March 22nd, 2020. The decreasing of the gejala corona related search occurred since March 22nd, although newly confirmed COVID-19 cases began to rise in Indonesia. Analysis using time-lagged correlation had a significant result between newly confirmed COVID-19 cases and the Google Trends data for Corona, Gejala Corona, Mencegah Corona and Obat Corona with p-value ≤0.05 and R-value range from 0.2-0.5 as shown in Table 3. Mencegah Corona related searches found to be the highest R-value in the overall time period (R value=0.46-0.58), while the day a number of newly Covid-19 cases met GT RSV had the weakest R-value range from 0.2-0.3. In addition, three days before until the onset of newly confirmed COVID-19 cases in Indonesia showed the highest R-value.
Analysis and potential use of google trends as a monitoring tool for risk... (Nindya Mahfuza)
The search pattern around COVID-19 terminology in Indonesia was very dependent on the media coverage and regulation by the Government; it's not due to the behavior of people who want to find out about the COVID-19.

3.1. Terminology

The keywords search pattern around COVID-19 terminology in Indonesia presented in Figure 1. The results yielded that the terminology used to define COVID-19 in Indonesia was 'corona' compared to worldwide result that used 'coronavirus' as a commonly used terminology as shown in Figure 2, as well as a search in Malaysia as the neighboring country of Indonesia which was found using the terminology ‘COVID’ and ‘coronavirus’ [10, 11]. This difference may occur because Indonesia has its distinctions in determining terminology.

There were five peaks for corona in Indonesia. The first peak happened on January 27th, 2020, when Ministry of Health of Indonesia release COVID-19 distribution data to the public that showed 15 patients in observation and 3 monitored, alongside the announcement by Chinese Government that the Lunar New Year's holiday would be extended until February 2nd, 2020 [12]. This increase did not last long and immediately decrease, although on January 30th, 2020, WHO declared COVID-19 as a global health emergency [13]. The results also differed from the search trends in Malaysia that found the first peak on January 27th which remained constant and did not have a significant decline [9]. Meanwhile, searches related to COVID-19 in Taiwan remained low after the first case was detected at Wuhan, China and experienced an increase after the first import case in Taiwan on 21 January 2020. And experienced a peak of COVID-19 related search occurred on 30 January 2020 in which WHO reported as PHEIC [14, 15].

The second peak happened when the first case of the COVID-19 infection in Indonesia was confirmed on March 2nd, 2020 [16]. The increase only lasted for a day, then immediately went down. The third peak was found on March 16th, 2020, chronologically, the National Disaster Management Agency (National Disaster Management Authority/Badan Nasional Penanggulangan Bencana) declared the extension of the disaster emergency status of COVID-19 until May 29th, 2020 and the Government's policy of Kemendikbud for online study and work from home, then the policies followed by other agencies of Government [17]. The search remained constant until the fourth peak appeared on March 22nd, 2020, when the Government announced positive cases of COVID-19 rapidly reached 514 people. The figure then decreased and again increased on April 4th, 2020, when media reported the number of COVID-19 cases in Indonesia has exceeded 2000 [2].

The search pattern of COVID-19 which represented using the ‘corona’ keyword in Indonesia shows that the search in the community was very dependent on the media coverage and regulation by the Government; it’s not due to the behavior of people who want to find out about the COVID-19.
It was evident from the number of peaks that skyrockets and instantly decreases, while the results of trends around the world were relatively constant with less peak fluctuation with less peak fluctuation [10]. There was some important world-related momentum for COVID-19 that did not increase the number of searches in Indonesia, such as on December 31st, 2020, when China reported to WHO there was an outbreak of pneumonia in Wuhan, on January 30th, 2020, when WHO declared a global health emergency, and the same result was also found on March 11th, 2020, when WHO declared COVID-19 as a pandemic, but the graphic of the search-related to COVID or Corona is still relatively low [18].

3.2. Correlation

People’s interest in COVID-19 in Indonesia mostly about symptoms of corona (gejala corona) followed by corona drug (obat corona) and preventing corona (mencegah corona). These results may indicate characteristics of information-seeking behavior in Indonesia that closely related to the diagnosis then preventive and curative. These results differ from those found in the United States, where in the United States the most relevant coronavirus keyword is 'coronavirus prevention' [19].

Information-seeking behavior is influenced by several variables [20]. An additional factor affecting information-seeking behavior is the keyword suggestion provided by Google. According to Weng and Sun [21], the suggested keyword by Google is based on the results of the last query from the ranking of content. Therefore, it can be concluded that conditions, distributions and dynamic factors affecting information-seeking behavior can vary in each country and change from time to time. However, the identification of factors affecting information-seeking behavior is needed to support the use of Google Trends as a monitoring risk communication tool in Indonesia.

The decreasing of the gejala corona (corona symptoms) related search occurred since March 22nd, although newly confirmed COVID-19 cases began to rise in Indonesia. According to Kata [22], information that people get will likely affect the medical decision they make. Supported this condition, on March 22nd, 2020, the Government stated that the community finally has a better understanding of the importance of social distancing [23]. Information that people have about corona symptom has already enough, so people’s interest begins to change. The Rising of COVID-19 related information using other online communication media such as Whatsapp, Line, Instagram and Facebook also suspected to be the cause of the condition.

Instagram was an example has been examined in the past, able to assess the public sentiment against the Zika virus during the outbreak [24]. Supported that, Instagram can be an ideal platform to maintain interactive communication between health organizations and community targets during an Ebola outbreak [25]. The increase of health organizations activity in Indonesia and the world, such as Kemenkes RI, WHO, CDC and others on the Instagram platform to educate the public about COVID-19 and answer public questions could be the reason for the decline of gejala corona related searches on Google, although newly confirmed COVID-19 cases continue to increase.

Unlike 'gejala (symptom) corona' keyword that suffered a sharp decline, people interest about 'mencegah (preventing) corona' remained constant, while the number of newly confirmed COVID-19 cases in Indonesia began to increase. This indicates that people’s interest in COVID-19 related search in Indonesia began to change, showed with increased interest in prevention against COVID-19 searches. Obat corona related search in Indonesia was experiencing its peak on March 20th, 2020; this situation coincided with the President’s announcement that Indonesia has been ordering 2 million of the drug (favipiravir) and three million of chloroquine [26]. A lot of media coverage on this announcement contributed to the spike that happened. However, this search gradually declined.

Validation using Pearson correlation shows a significant correlation (p-value ≤ 0.05) between GT data of corona, gejala corona, mencegah corona, and obat corona and newly confirmed COVID-19 cases in Indonesia as shown in Table 3. These results indicate that the increase of search activity through Google regarding gejala corona, mencegah corona, and obat corona occurred three days before until the onset of newly confirmed COVID-19 cases in Indonesia.

The survey released by the Indonesian internet service provider association (APJII) about penetration and behavior of Internet users in 2018 revealed that Indonesian Internet users increased by 10.12% in 2018 compared to the previous year [27]. In total, internet users reached 171.17 million users from 264.16 million total populations. This high internet penetration rate in Indonesia can be an opportunity to determine the best risk communication during this pandemic.

Risk communication purposes are to share vital information about saving lives, maintaining health, reducing losses, giving respect, and/or changing behavior. The proper risk communication during a pandemic is important to reduce the impact, hence, WHO stated that the result of risk communication should be measured [3]. Without monitoring and assessment of changes in knowledge, habits, and behavior, risk communication becomes meaningless and can not help to control and regulate the public health.
emergency. Unmonitored results can lead to ineffective results and make wrong thoughts about who should be responsible for the public response. Surveillance research using Google Trends that compares between GT data and datasets has been extensively performed and obtained strong association results [7, 9, 28, 29]. In line with this study, GT can be used as a tool to monitor public unrest related to the COVID-19 in Taiwan within 1-3 days before the increase of newly confirmed COVID-19 [9]. A study found a strong correlation between searches on HIV and incidence of HIV cases based on CDC, they can also create a model based on a search pattern from 2007 to 2008, to accurately predict HIV incidence in 2009-2010 [30]. Supported this study, GT data has a linear pattern and significantly correlated with dengue fever disease (DBD) reports in Indonesia [20].

These results concluded that GT can be used as a tool to monitor public restlessness over the COVID-19 pandemic in Indonesia within 1-3 days before the increasing number of cases occurred. Improvement in analyzing public restlessness about the major outbreak can help maximize the risk communication conducted by the Government. Its purpose to change the community behavior, through providing information on the importance of self-protection, introducing the symptoms to the general public, especially the high-risk population [7]. So it can be expected to reduce the impact of the pandemic. Therefore, the opportunity to use Google Trends as an early warning system and a public reaction monitoring tool that used to assist the traditional surveillance system and increase public health responses to the COVID-19 pandemic has increased [7, 9, 28, 29].

4. CONCLUSION

The search pattern for COVID-19 in Indonesia has its features. The terminology for search-related COVID-19 in Indonesia is using ‘corona’ keyword. Based on Google Trends data, corona-related searches in Indonesia are extremely fluctuating and depend on news provided by the Government. In the early declaration of the pandemic in Indonesia, the corona-related search has not shown any improvement. This result shows that risk communication in Indonesia has not been applied well. Along with the increasing cases of COVID-19, the response of public concerns also began to be seen, supported with significant results from the time-lagged correlation between ‘corona, gejala corona (corona symptoms), mencegah corona (preventing corona), and obat corona (corona drugs)’ compared to newly confirmed COVID-19 cases in Indonesia. Therefore, proper information through risk communication that provided during the pandemic using Google Trends data might maximize the improvement of risk communication; also Google Trends can potentially be used as a tool to monitor public restlessness during the COVID-19 pandemic in Indonesia by government.

REFERENCES

Int. J. Public Health Sci ISSN: 2252-8806


Analysis and potential use of google trends as a monitoring tool for risk... (Nindya Mahfuza)