Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The impact of the pandemic declaration on public awareness and behavior: Focusing on COVID-19 google searches

Seung-Pyo Jun \textsuperscript{a,b}, Hyoung Sun Yoo \textsuperscript{b}, Jae-Seong Lee \textsuperscript{a}

\textsuperscript{a} Data Analysis Platform Center, Korea Institute of Science and Technology Information and Science & Technology Management Policy, University of Science & Technology (UST), 66, Hoegi-ro, Dongdaemun-gu, Seoul 130-741, Korea
\textsuperscript{b} Korea Institute of Science and Technology Information and Science & Technology Management Policy, University of Science & Technology (UST), 66, Hoegi-ro, Dongdaemun-gu, Seoul 130-741, Korea

\textbf{A R T I C L E   I N F O}

\textbf{Keywords:}
Coronavirus
COVID-19
Public awareness
Google Trends
Relative search volume
Pandemic declaration

\textbf{A B S T R A C T}

The unprecedented outbreaks of epidemics such as the coronavirus has caused major socio-economic changes. To analyze public risk awareness and behavior in response to the outbreak of epidemic diseases, this study focuses on RSV (Relative Search Volume) provided by Google Trends. This study uses the social big data provided by Google RSV to investigate how the WHO’s pandemic declaration affected public awareness and behavior. 37 OECD countries were analyzed and clustered according to the degree of reaction to the declaration, and the United States, France and Germany were selected for comparative study. The results of this study statistically confirmed that the pandemic declaration increased public awareness and had the effect of increasing searches for information on COVID-19 by more than 20%. In addition, this rapid rise in RSV also reflected interest in the COVID-19 test and had the effect of inducing individuals to be tested, which helped identify new cases. The significance of this study is that it provided the theoretical foundation for using RSV and its implications to understand and strategically utilize public awareness and behavior in situations where the WHO and governments must launch policies in response to the outbreak of new infectious diseases such as COVID-19.

\textbf{1. Introduction}

This study investigates how the declaration of the coronavirus (COVID-19) pandemic, which was the third time the World Health Organization (WHO) declared a pandemic since its establishment in 1948 (Broberg, 2020), affected the public. Specifically, this study attempts to empirically demonstrate the ongoing human behavioral changes and impact on public awareness of COVID-19 by tracking Google search data before and after the WHO’s pandemic declaration. Here, “public awareness” refers to the public’s level of understanding about COVID-19 (Capano et al., 2020). Google search data can objectively measure the public’s information search behavior (Jun, 2012a; Jun et al., 2016), and this study utilizes the relative search volume (hereinafter referred to as RSV) provided by Google Trends (Do et al., 2015; Husnayain et al., 2020). This empirical analysis is critically important because changes in public awareness regarding a global pandemic is one of the key factors in diagnosing and preventing new cases and strengthening group immunity (Do et al., 2015; Hu et al., 2020; Randolph and Barreiro, 2020). Therefore, this study analyzes how the WHO’s pandemic declaration, an important event, affected public risk awareness (or perceptions) and behavior in response to the outbreak of this epidemic disease.

The key question investigated in this study is, “How did the WHO’s pandemic declaration affect public awareness and behavior at the beginning of the global COVID-19 outbreak?” We aim to demonstrate how interventions such as the WHO’s pandemic declaration can play an important role in strengthening public awareness as well as the responses of policy authorities in each country, by analyzing public reactions as reflected in information search activities in response to the WHO’s pandemic statement. In addition, this study will illuminate how these phenomena can be observed, how their results or effects can be measured, and furthermore, how these observations and measurements can be utilized in the future.

To figure out how we ought to interpret and utilize RSV changes in the unprecedented situation of a new pandemic such as COVID-19, we need to understand the determinants and impacts of RSV changes. We need to know the determinants of RSV changes to understand the significance of RSV in the process of monitoring RSV, and we also need to understand these determinants to make decisions based on the...
monitored RSV changes. In addition, the analysis of determinants will also indicate which variables should serve as the basis for clustering countries for hypothesis testing. To analyze the RSV determinants, this study cross-sectionally analyzed changes in RSV in 37 OECD countries, changes which occurred before and after the declaration. To analyze the impact of RSV, we analyzed and compared the temporal changes of RSV from a longitudinal perspective over 100 days in the cases of the United States, Germany, and France, which were countries that were identified as belonging to different clusters in our determinants analysis. We thereby determined how the rise of RSV varied due to the declaration, how the increase persisted, and how it was correlated to new tests and new (confirmed) cases of COVID-19.

This study assumed that the policy changes implemented in each country immediately after the declaration are within the scope of the same event effect. Section 2 reviews preceding studies on infectious diseases that used RSV, including studies on policy changes in each country immediately after the declaration, and reviews other recent studies related to COVID-19 that were conducted based on social data. Our research model was constructed on the premise that changes in awareness will manifest as changes in information search behavior, from the perspective of human behavior; relevant explanations on this point are provided in Section 3 along with the explanation of our research methodology. Section 3 also explains the data used in this study. Section 4 presents our research results with explanations of the findings of our analyses of RSV determinants and RSV effects. Section 5 concludes with a discussion of the results of this research and the policy implications for utilizing the results.

2. Preceding research and recent policy trends

2.1. Infectious diseases and RSV

Risk awareness among the public is critically important to prevent the spread of infectious diseases (Guo et al., 2015). It is possible to significantly reduce the spread of infectious diseases by pre-emptively disseminating knowledge about how widely infectious diseases are spreading and how severe the outbreaks will be and educating the public on precautionary measures (Funk et al., 2009; Wang et al., 2019). In addition, to effectively respond to the spread of infectious diseases, it is important for people to recognize the general symptoms of infection, so that they may determine whether or not they are infected by comparing their own symptoms with those of people around them, and then take appropriate measures accordingly (Zang, 2018). In the past, people mostly relied on mass media and word of mouth to obtain information about an infectious disease and how to prevent it and respond to it (Wu et al., 2012). However, with the availability of the internet and the widespread use of search engines, people began to seek information more actively, and this effort is reflected in the data of their search activity. There have been many studies of behavioral analysis or forecasting using RSV in various fields, not only regarding infectious diseases, and in many such studies, RSV was understood through the lens of the human behavior processes described in theories such as the five stages of the consumer buying process or the innovation adoption model (Kotler et al., 2014; Rogers, 2003). According to these theories, problem awareness leads to information searching and the seeking of alternatives for problem solving (Jun et al., 2014). In the same vein, RSV regarding infectious diseases has been found to be useful for quantifying the level of people’s awareness of infectious diseases (Boehm et al., 2019). Through specific keyword analysis, it is possible to distinguish, for example, whether people want to know “what” infectious disease is spreading and “what” kinds of symptoms indicate infection, or whether they are seeking information on “how” to prevent and treat it (Cairo, 2020). Because Google has a dominant share (92%) of the global search engine market (Statcounter, 2020), the RSV of Google searches has been recognized to be an especially effective means of gauging people’s perception of infectious diseases (Do et al., 2015; Hu et al., 2020).

In this context, researchers have steadily used RSV to observe and monitor infectious diseases, identify awareness levels, and predict spread (Nuti et al., 2014). The first leading example is the study by Ginsberg et al. (2009), who presented a model that predicts current flu levels using Google Trends, the RSV provided by Google, and demonstrated that it can predict the outbreak of influenza one to two weeks earlier than the Center for Disease Control (CDC). Pellet et al. (2009) analyzed the relationship between RSV and the incidences of three other infectious diseases, namely ILI, gastroenteritis, and chickenpox. Based on their results, they reported that RSV can be used broadly to monitor diseases other than influenza and that even applying one or two well-chosen keywords can be highly effective. In addition, it has been reported that RSV can be effectively used to monitor the spread of various other infectious diseases such as Lyme disease, dengue fever, and Zika virus (Althouse et al., 2011; Seifert et al., 2010; Strauss et al., 2020). After the release of the Google Flu Trends (GFT) service using Google’s RSV, the accuracy of these predictions was recognized in many countries (Cook et al., 2011; Dugas et al., 2012). However, researchers have also steadily pointed out problems in the accuracy of GFT and some of the fundamental limitations of predicting infectious diseases using RSV (Butler, 2013; Lazer et al., 2014; Olson et al., 2013; Pollet et al., 2016). Butler (2013) reported that the GFT algorithm greatly overestimated the peak flu level in 2013. Butler argued that because RSV can be affected by external effects such as the mass media, and because it is difficult to identify the reasons motivating searches, there may be limitations in using RSV to predict the level of infectious diseases. Lazer et al. (2014) also cautioned against hubris, arguing that while social big data such as RSV can be a useful tool for understanding people’s behavioral patterns, the accuracy of predictions based on such data is limited, and therefore such data should be used to supplement rather than replace conventional methods. The GFT service was terminated in 2015 due to various issues, but there have been continued efforts to improve the performance of nowcasting and forecasting of infectious diseases using RSV (Lampos et al., 2015; Preis and Moat, 2014; Santillana et al., 2014). Yang et al. (2015) used Google Trends to propose ARGO (AutoRegression with Google search data), an influenza tracking model with more accurate prediction performance than GFT. Furthermore, more recent studies have evaluated that RSV can provide higher accuracy and predictive power when used in combination with conventional monitoring methodology, stronger compared to when each is applied individually (Kandula and Shaman, 2019). Althouse et al. (2019) used Google Trends data to confirm a strong correlation between search traffic and the number of hospitalizations. Based on these results, the authors argued that Google Search data can be useful for decision-making on medical and public health issues because it has the advantage of being updated quickly without additional costs.

Meanwhile, a number of studies have been published recently on the relationship between RSV and COVID-19, which has spread around the globe; these will be discussed in detail in the next section.

2.2. COVID-19 and social big data

On the subject of COVID-19, many studies have used sources of social data such as news, Twitter, and Facebook, or attempted quantitative approaches using RSV, as we do in this study. The main areas of research conducted in this manner can broadly divided into two. First, there are studies that tracked changes in the public’s psychology regarding the epidemic to determine the country’s general response to infectious diseases. The other remaining studies can be sub-divided into studies empirically analyzing unconfirmed new symptoms, the number of new cases, or the number of hospitalizations. The contents of these preceding studies are reviewed in the following.

The proliferation of an infection disease such as COVID-19 also has a significant impact on public psychology (Katafuchi et al., 2020). First, studies on psychological changes among the public using social big data...
include the following examples. Li et al. (2020b) studied the changes in the results of sentiment analyses of posts written by Weibo users before and after the WHO pandemic declaration. The results demonstrated that after the WHO’s pandemic declaration, indicators of anxiety, depression, and indignation increased significantly. In addition, the study found that social risk judgment was significantly increased, and life satisfaction was significantly decreased. Hamidein et al. (2020) conducted a survey of participants who regularly checked COVID-19-related news and found that they generally felt negative emotions in response to the news but were also deriving problem-solving strategies through the news, identifying what can be done at an individual level to prevent the spread of the virus. Brodeur et al. (2020) focused on the lockdown implemented by many governments in response to COVID-19. The authors found that lockdowns may help to contain the spread of the virus but warned that it may result in substantial damage to the population’s well-being. To assess the level of such damage, this study used Google Trends data to analyze the causal effects, tracking how topic search terms related to well-being changed before and after the lockdown policy. Based on this analysis, the study evaluated the impact of governments’ lockdown policies on the public’s mental health. Chen et al. (2020) evaluated public perception of COVID-19 through Weibo, a Chinese social media platform. This study underscored the importance of how experiences of outbreaks of similar diseases in the past affect the public’s perception of the outbreak of a new infectious disease.

Meanwhile, the following studies attempted to provide empirical analyses of major issues related to COVID-19 using social data such as RSV as quantitative data. Walker et al. (2020) noted that while initial reports describing COVID-19 symptoms mainly pointed to cough, breathlessness, and fever, there were as yet unverified reports suggesting that anosmia may also be a manifestation. Their study used Google Trends to investigate the presence of this unverified symptom, by tracking whether there was an actual surge in individuals searching for information related to smell loss during the COVID-19 epidemic in several major countries. The study found strong correlations between loss of smell and increases of daily COVID-19 cases and deaths in all of the major countries. The study by Hu et al. (2020) adopted a similar perspective and used Google Trends to investigate the global public perception of COVID-19. The study analyzed the period of two months before and after January 30, the date on which the WHO announced a Public Health Emergency of International Concern (PHEIC), and claimed that the first peak was observed immediately after the WHO’s PHEIC announcement. This study examined the correlation between RSV and daily confirmed patients and found statistically significant positive correlations in many countries. The authors pointed out that the reaction time was different in each country and the overall duration of public attention was short, and argued that efforts should be made to strengthen public awareness on COVID-19 at the national level and to reinforce public vigilance and sensitivity to COVID-19 (Hu et al., 2020).

There have been various studies that demonstrated that RSV has a very significant correlation with new cases in individual countries. Husnayain et al. (2020) reported that moderate to strong correlations between Google relative search volumes (RSVs) and COVID-19 cases were found in Taipei, and provided studies using specific search terms such as face mask, handwashing, etc. Ayyoubzadeh et al. (2020) analyzed data in Iran to show that search terms such as handwashing and hand sanitizer can be significant variables in predicting COVID-19 infections. Li et al. (2020a) analyzed internet searches conducted in Baidu and Weibo as well as Google, and argued that the COVID-19 outbreak occurred 10–14 days earlier than the peak of daily incidences in China, demonstrating the internet search data can have effective surveillance capabilities for tracking the outbreak of a new disease.

Cairo (2020) analyzed search patterns found in RSV and found that in the case of the United States, after the first confirmed case in the country, there was a brief increase in searches, most of which asked “what is/are,” but after February 29, when the first deaths occurred and the number of confirmed cases in other countries increased significantly, the searches surged significantly again. Specifically, prior to March, most searches were queries for general information such as “What is coronavirus?” but beginning in March, especially after the actions of the WHO and the United States government, there was a surge in searches for “What are symptoms of coronavirus?” or “how to” information.

As summarized above, most studies analyzing social data related to COVID-19 focused on changes in the public’s emotional states such as anxiety, sadness, or stress, in response to governments’ lockdown policies to stem the spread of the infectious disease. While the studies reviewed above also researched the effects of the pandemic declaration, our study differs in that we focused on changes in the public’s search patterns in response to the outbreak of disease and their subsequent behavior instead of focusing on changes in the perceived emotional state of the general public. What distinguishes this study is that we explain the serial process leading from risk awareness to information searching, testing, and confirmed cases through the analysis of changes in RSV.

2.3. The intervention of the WHO and policy changes in each country

The WHO proclaimed a Public Health Emergency of International Concern (PHEIC) on January 30, and eventually declared the situation a pandemic on March 11, which led many countries to adopt responsive policy measures. As mentioned above, in this study we regard the policies implemented immediately after the WHO pandemic declaration as the effect of the same event and therefore we found it necessary to understand the policy changes that occurred immediately after the WHO pandemic declaration. We briefly examined the policy changes in the three countries we selected as case studies based on our longitudinal study.

On March 11, the WHO declared the outbreak to be a pandemic. By this time, the virus had spread to 110 countries (The local.de, 2020) and all continents except Antarctica (Nakamura and Managi, 2020; WHO, 2020). The WHO’s definition of a pandemic “mixed severity and spread,” reported Vox, and it held off calling the outbreak a pandemic because many countries at the time were reporting no spread or low spread (Piper, 2020). In the United States, on March 11, when the WHO declared COVID-19 to be a global health pandemic, later that same day President Trump announced new restrictions on travel from Europe (Goodman, 2020). In a prime time, Oval Office address to the nation, President Trump announced new travel restrictions to and from 26 European countries, said that the United States had taken early intense action, and commended the “dramatically fewer cases of the virus in the United States than are now present in Europe.” Travel restrictions went into effect on March 13 for 26 European countries in the Schengen Area and similar restrictions entered into effect on March 16 for the United Kingdom and Ireland (Woodward, 2020). On the same day, President Trump declared that the coronavirus pandemic to be a national emergency. He also stated: “I don’t take responsibility at all,” in response to a question about the lack of available tests. Also, on March 16, the White House advised against any gatherings of more than ten people (Liptak, 2020).

Meanwhile in Germany, the citizens of Germany were shocked when President Trump of the United States announced on March 12 the 30-day travel ban for foreigners who travelled from Schengen area states, including Germany. German politicians criticized the United States for adopting this measure without prior agreement, and complained of the exclusion of the United Kingdom from the ban (Vitzthum, 2020). On March 13, 14 of the 16 federal states of Germany decided to close all kindergartens, schools and Kitas (daycare centers) up to April 6 (The local.de, 2020). Germany’s neighbors, the Czech Republic, Poland and Denmark, closed their borders (BBC news, 2020a). On March 22, the government and the federal states prohibited gatherings of more than two persons for a minimum of two weeks and agreed to impose a social distancing rule of 1.5 meters between individuals in public spaces with
the exception of family members, partners or other cohabitants. Service establishments such as restaurants and hair salons were closed on the same date (Ohms, 2020).

In France, on March 12, President Emmanuel Macron announced on public television that all schools and universities would close on Monday, March 16, until further notice (Cuthbertson, 2020b). The next day, Prime Minister Édouard Philippe banned gatherings of more than 100 people, not including public transport. The following day, the prime minister ordered the closure of all non-essential public places, including restaurants, cafes, cinemas, and nightclubs, effective from that midnight (BBC_News, 2020b). On March 16, President Macron declared a compulsory lock-down in homes to go into effect at noon on March 17, for a duration of 15 days. This lock-down measure was extended twice and ended only on May 11 (Cuthbertson, 2020a).

3. Research design

3.1. Research model and method

Synthesizing the preceding studies, we can conclude that the WHO’s pandemic declaration increased the public’s social risk awareness and judgment, and using the information they already had or newly acquired (for example, through searches), the public adopted problem-solving strategies that could be performed at an individual level (Hamidein et al., 2020; Husnayain et al., 2020; Li et al., 2020b). It has also been shown that the above process is closely related to public awareness (Cairo, 2020; Chen et al., 2020; Hu et al., 2020). The studies also demonstrated that search data can help identify new symptoms or may have high correlation with the number of new cases, indicating that such data is useful for quantitatively analyzing major issues related to COVID-19 (Ayyoubzadeh et al., 2020; Li et al., 2020a). We also confirmed that this series of processes is in line with the human behavior process described in the consumer buying process or innovation adoption model (Jun et al., 2014; Koller et al., 2014; Rogers, 2003).

Based on these existing studies and theories, we judged that once the WHO pandemic declaration raised public awareness, this would arouse public interest and increase RSV as people seek information to understand COVID-19 and find countermeasures. In addition, we hypothesized that the increased RSV would contribute to identifying new confirmed cases by motivating people with symptoms to take tests. Fig. 1 presents our research methodology in more detail.

The two hypotheses were both posited from the perspective of the impact of the COVID-19 pandemic declaration on awareness (or RSV). Hypothesis 1 is about the effects of the WHO COVID-19 pandemic declaration on awareness, as explained in Fig. 1, and about the changes in behavior resulting from this shift in awareness. To test this hypothesis, we must first conduct a determinant factor analysis to identify the factors that can affect changes in RSV. This is partly because of the practical need to understand these determinant factors to monitor and strategically utilize RSV, but also because the COVID-19 declaration may be a determinant factor for awareness, or RSV, and each country exhibited varying RSV responses to the COVID-19 events (Hu et al., 2020). To select the countries in which to test our hypotheses regarding the impacts of WHO pandemic declaration, we must take the determinants into consideration. Therefore, this study was divided into two major areas of research, centering on the information search (RSV) shown in Fig. 1, analyzing the determinants of RSV changes and the impacts of RSV. Fig. 2 presents our research methodology in more detail.

In the RSV determinants analysis in Fig. 2, which was designed to investigate the question “What determinants affected RSV other than the WHO declaration?” we performed case analyses while controlling for the effect of the pandemic declaration intervention. We looked for determinants that affected the rate of change in RSV in the event week (from March 11 to March 18) compared to the previous week (from March 4 to March 10) in OECD countries.1

Based on previous comparative studies of countries regarding COVID-19, the median age, share of the elderly population, GDP, and hospital beds were considered as control variables for identifying the determinants (Chakraborty and Ghosh, 2020; Wang et al., 2020; Wu et al., 2020). Total COVID-19 tests and new cases were considered as COVID-19 situation variables. Here, the new cases variable was also analyzed as the rate of change in the event week compared to the previous week. As for the total number of tests, we analyzed the cumulative data as a country-specific constant value (square root), like a control variable.2 Certain limitations in the data on COVID-19 tests forced us to examine the two situational variables using different types of data, respectively in the form of the rate of change and a constant value. On March 11, when the WHO COVID-19 declaration occurred, there was a relatively large number of countries in which COVID-19 incidence was still in the early stages, and therefore daily or weekly test statistics were available in many cases. For this reason, for the variable of the tests by country, we used data on cumulative tests up to March 18, which was more easily available, rather than using the rate of change. Also, for new cases, the ratio was directly analyzed, but for cumulative tests, we converted the data to the square root. This was because as we analyzed the relatively few cases of the 37 OECD countries, we found that the distribution of the cumulative test values deviated greatly from normality. We used regression analysis as our research method, as shown in Fig. 2, and for our main method of inference, we adopted the best subset selection method, which is superior even to conventional regression analysis based on the method of forward or backward stepwise selection (James et al., 2013). The problem with the conventional method of forward stepwise selection is that it fails to exclude already included variables, while in backward stepwise selection, the problem is that variables that have already been excluded are not included again. The mixed (or hybrid) method solves some of the problems of forward and backward stepwise selection but fails to completely eradicate the same problems (James et al., 2013). By contrast, the best subset selection method fits all possible subsets and identifies the best subset by comparing criteria such as Cp, AICC, and BIC. To overcome the limitations of regression analysis based on the existing least squares method, methods such as Lasso based on statistical learning have been proposed (Hastie et al., 2017; James et al., 2013). We also conducted Lasso analysis for this study, and the results obtained from the Lasso analysis is presented below briefly, for the sake of comparison. In this study, the results of the Best Subset Selection and Lasso methods were similar, but in the Lasso method all variables were required to have no missing values and this restriction made the method applicable to fewer cases, and furthermore, the data available was somewhat insufficient for performing statistical learning. Based on these considerations, we chose to proceed with our explanation based mainly on the results from the Best

---

1 We analyzed changes in weekly data for a week (March 11 – 18) starting from March 11 when the event occurred. To offset the seasonal changes of daily data.

2 In the case of the cumulative number of tests, we used the data as of March 18 because there were many cases of countries where data collection was omitted before March 11. For Australia, Chile, Denmark, France, Germany, Iceland, Italy, Luxembourg, Netherlands, Spain and Sweden, we used values estimated by interpolation or extrapolation.
Subset Selection method.

In our analysis of determinants, we next performed clustering analysis to find how the differences across countries due to determinants can be clustered. Specifically, we applied the DBSCAN method. DBSCAN refers to a density-based clustering method that performs clustering by measuring the density of an area. Other examples of such density-based methods are OPTICS, ADWICE and BIRCH (Ankerst et al., 1999; Buibek and Nadim-Tehrani, 2004; Burbeck and Nadim-Tehrani, 2007; Thang and Kim, 2011; Zhang et al., 1996). Among these density-based methods, DBSCAN, a method developed by Ester et al. (1996), is one of the most common clustering algorithms and is the most frequently cited method in scientific literature (Chakraborty and Nagwani, 2014).

Through the above determinants analysis, we first reduced the range of variables to be considered in the analysis of the impact of the declaration, as shown in Fig. 2. Furthermore, through cluster analysis, we narrowed the target of analysis to be used to test our hypotheses.

As shown in Fig. 2, the analysis of the impacts of the declaration and RSV change is the second area of our research, aiming to test our established hypotheses. First, to test Hypothesis 1, we investigated the question “To what degree did the WHO pandemic declaration affect RSV?” According to the differences by country that appeared in the results of the first area of our analysis, we selected three countries to be representative of the clusters (United States, Germany and France) and analyzed them comparatively. We analyzed whether the change in the event week, more precisely, the change in RSV on March 12, was statistically different when viewed in a time series analysis; that is, we analyzed whether the effects of the intervention were observable. It was possible that the RSV rise in the event week may simply be the extension of an existing trend and may not be a special increase in response to the pandemic declaration. Time series data often show rapid changes when affected by specific external events such as oil shocks, holidays, and policy changes. We refer to these external events as interventions, and the method of evaluating the effects of these external events is called intervention analysis (Box and Tiao, 1975; Montgomery and Weatherby, 1988; Ray et al., 2017).

To test the second hypothesis, namely, to answer the question “Does an increase in RSV result in an increase in the detection of confirmed cases?” we performed the second analysis on the second subject, as described in Fig. 2. We analyzed the correlation between RSV and new cases, statistically analyzing the relationship using the methods of cross-correlation, Granger causality, and cointegration tests for analyzing VAR (Vector Autoregression) or VECM (Vector Error Correction Models) (Gharehgozli et al., 2020; Jun and Park, 2016; Jun et al., 2017). In particular, we analyzed changes in COVID-19 testing following the increase in RSV to explain the correlation between RSV and new cases. The VAR model we used here is a multivariate time series model that combines the features of time series analysis and regression analysis to estimate the correlation and causal relationship between variables (Engle and Granger, 1987; Engle and Yoo, 1987). In the cointegration test between variables, if there is a cointegration relationship between variables, VECM can be used in consideration of the long-term equilibrium. Therefore, instead of examining the relationship between variables with a VAR model, we only conducted cointegration analysis (Johansen, 1988; Jun et al., 2016).

3.2. Data collection and cases

To evaluate these research hypotheses, we collected not only data from Google but also search statistics from the ECDPC (European Centre for Disease Prevention and Control) and the OECD, and data on new cases and new deaths and other information required for control from 37 OECD member countries. The time period subject to analysis was the period from February 1, immediately following the WHO’s declaration.

3 Although the WHO’s announcement was made on March 11, we analyzed the RSV as of March 12. Because the standard time varies by country, and we also needed to consider the time of the release of WHO’s announcement, we judged that March 12 would be the date that best reflects the public’s response. Furthermore, we also considered that there were many countries in which the highest value of RSV (100) was reached on March 12, or in other words, in which the RSV showed a rapid increase, as shown in Table 2.
Table 1
Key research variables.

| Variable          | Description                                    | Source                                                                 | Reference                      |
|-------------------|------------------------------------------------|------------------------------------------------------------------------|--------------------------------|
| RSV               | Relative Search Volume                        | Google Trends, Coronavirus Search Trends                               | Walker et al., (2020)          |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             | Huynhavan et al., (2020)       |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             | Huynhavan et al., (2020); Hu et al., (2020) |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             | Hu et al., (2020)             |
|                   |                                                | Althouse et al., (2019)                                               |                                |
| New cases         | New daily confirmed cases of COVID-19 per 1,000,000 people | Althouse et al., (2019)                                               |                                |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             |                                |
|                   |                                                | National government reports                                          |                                |
| New deaths        | New daily deaths attributed to COVID-19 per 1,000,000 people | Althouse et al., (2019)                                               |                                |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             |                                |
|                   |                                                | National government reports                                          |                                |
| New tests         | New daily tests for COVID-19 per 1,000 people  | Althouse et al., (2019)                                               |                                |
|                   |                                                | European Centre for Disease Prevention and Control (ECDPC)             |                                |
|                   |                                                | National government reports                                          |                                |
| Total tests       | Total tests for COVID-19 per 1,000 people     | European Centre for Disease Prevention and Control (ECDPC)             |                                |
|                   |                                                | National government reports                                          |                                |
| Population        | Population in 2020                            | United Nations (UN), Department of Economic and Social Affairs, Population Division, World Population Prospects: The 2019 Revision | Wu et al., (2020)             |
| Median age        | Median age of the population, UN projection for 2020 | UN Population Division, World Population Prospects, 2017 Revision | Chakraborty and Ghosh, (2020) |
| Aged 65 older     | Share of the population that is 65 years old and older, most recent year available | World Bank – World Development Indicators, based on age/sex distributions of United Nations Population Division’s World Population Prospects: 2017 Revision | Wang et al., (2020)           |
| GPD (GDP per capita) | Gross domestic product at purchasing power parity (constant 2011 international dollars), most recent year available | World Bank – World Development Indicators, source from World Bank, International Comparison Program database | Wang et al., (2020)           |
| Hospital beds     | Hospital beds per 100,000 people, most recent year available since 2010 | OECD, Eurostat, World Bank, national government records and other sources | Chakraborty and Ghosh, (2020) |

Source: OWID (2020) modified.

that COVID-19’s risk assessment at the global level was high, to May 11, which is two months from March 11, the date on which the WHO assessed COVID-19 to be a pandemic (WHO, 2020).

Table 1 describes the variables included in this study, our sources and related references (OWID, 2020). Here, we collected search information from 37 OECD countries via Google Trends, and since the data was collected by country, all search data was collected as daily relative search volume, that is, a normalized value based on the search volume on the day with the highest domestic search volume during the data collection period (with its value set as 100) (Jun et al., 2018). Note that Google Trends provides RSV for “coronavirus,” considered the thesaurus equivalent of COVID-19 (Google_Trends, 2020). In the case of non-English speaking countries, statistics were also processed with the word “coronavirus.”

The data on confirmed cases, deaths, and test recipients were also collected daily. As shown in Table 1, such data was collected by the ECDPC from 37 OECD countries. However, collection of data was omitted on some dates, in some countries. For comparison of all these variables across countries, the data on both new cases and deaths were standardized to represent incidence per 1 million people. The number tested was standardized to be expressed as the rate of tests per 1000 individuals. In the determinant factor analysis, due to the weak normality of the distribution, we converted the data to the square root values, as explained above.

Table 2 shows the changes in the RSV, the key variable of this study, from just before the event week to the mid-point, in the 37 OECD countries we analyzed. In addition, Table 2 shows the dates when the peaks of new cases and deaths existed during the analysis period. In Table 2, we see that RSV showed rapid changes in many countries around March 11 and March 12. 17 out of 37 countries (45.9%) also exhibited their peak immediately following the declaration (within 2 days) and 28 out of 37 countries (75.7%) reached their peak of RSV within a week. These results confirm the likelihood that the WHO’s pandemic declaration had a significant effect on the changes in RSV, while also indicating that there were differences among countries.

Table 2 also presents the duration of the analyzed RSV. Duration is defined as the period (in days) until the RSV was restored to that of March 10, a date just preceding the event week. Table 2 also shows significant differences by country: it is notable that Italy, Japan, and Korea, where a large number of confirmed cases had already been reported, experienced a very short duration. The durations presented in Table 2 had a very high correlation with the RSV increase rate in the event week (Pearson correlation coefficient 0.808, p-value<0.000), and this can be confirmed by referring to Fig. 3. We thus confirmed that the greater the increase in RSV, the longer the duration, and in consideration of multicollinearity, we did not add RSV duration as a variable in our analysis.

Finally, to control the differences among countries, we considered several variables described in Table 1. The variables included in this study were median age, the percentage of the elderly population (over 65 years old), GDP per capita, and hospital beds and the data were collected from sources including the UN, World Bank, and OECD. The definition of each variable is shown in Table 1. GDP per capita refers to national income per capita, but the normality of the distribution was weak, so the value was converted to log values for our analysis. These variables were assumed to be constant values within each country during the period of 100 days analyzed, as shown in Table 2.

As for the countries analyzed, as explained above, to analyze the first issue outlined in Fig. 2 we conducted a cross-sectional analysis of 37 OECD countries (including Colombia, since its participation was confirmed). To analyze the second issue explained in Fig. 2, we needed to specify selected countries to account for national differences through longitudinal analysis, and for this purpose we chose the United States, Germany and France. The cluster of the highest interest was the sensitive cluster, and to represent this cluster we selected the United States since it

---

4 In Table 1, Althouse et al. (2019) explains the correlation with hospital visits, rather than the correlation between RSV and tests.

5 We chose to express the number of tests per one thousand persons, in comparison with the scale used for the figures for new cases and new deaths, in order to maintain the same scale as provided in the original data and also in consideration of the number of digits in the significant figures.

6 In this study, the data on deaths were relatively strongly affected by additional factors such as national testing capability, medical infrastructure, and age distribution of the population and there were also differences among countries during the targeted research period. Therefore, these values were used only for reference purposes.
Table 2
Comparison of differences in major variables by country.

| Variable Country | Peak day RSV | New Cases 10-Mar | New Deaths 11-Mar | RSV (daily) 12-Mar | Median 13-Mar | Age 14-Mar | GDP 15-Mar | Duration | Median Age | Aged 65 older | Hospital Beds |
|------------------|--------------|------------------|------------------|-------------------|--------------|---------|---------|---------|-----------|-------------|---------------|
| Australia        | 22-Mar       | 22-Mar 4-Apr     |                  | 37 41 66 73 67    | 81           | 32      | 37.9   | 15.5    | 44,649    | 3.84        |
| Austria          | 23-Mar       | 23-Mar 4-Apr     |                  | 71 90 93 98 87    | 100          | 7       | 44.4   | 9.2     | 45,437    | 7.37        |
| Belgium          | 11-Apr       | 11-Apr 48 55     |                  | 100 90 82 88      | 14           | 14.8    | 18.6   | 42,659   | 5.64       |
| Canada           | 2-May        | 2-May 34 50      |                  | 100 99 79 82      | 32           | 32.2    | 14.1   | 17.0    | 44,018    | 2.50        |
| Chile            | 4-May        | 4-May 16 18      |                  | 32 41 52 83       | 49           | 35.4    | 11.1   | 22,767   | 2.11       |
| Colombia         | 6-May        | 6-May 5 31       |                  | 45 47 52 70       | 56           | 32.2    | 7.6    | 13,259   | 1.71       |
| Denmark          | 5-Apr        | 5-Apr 84 100     |                  | 96 74 65 57       | 3            | 42.3    | 19.7   | 46,683   | 2.50       |
| Estonia          | 27-Mar       | 27-Mar 38 49     |                  | 88 100 95 71      | 10           | 42.7    | 19.5   | 29,481   | 4.69       |
| Finland          | 22-Mar       | 22-Mar 41 48     |                  | 100 85 67 59      | 8            | 42.8    | 21.2   | 40,586   | 3.28       |
| France           | 1-Apr        | 1-Apr 44 47      |                  | 75 83 92 100      | 16           | 42.0    | 19.7   | 38,606   | 5.98       |
| Germany          | 16-Apr       | 16-Apr 56 62     |                  | 76 86 81 84       | 9            | 46.6    | 21.5   | 45,229   | 8.00       |
| Greece           | 5-Apr        | 5-Apr 71 76      |                  | 90 96 100 94      | 7            | 45.3    | 20.4   | 24,574   | 4.21       |
| Hungary          | 24-Apr       | 24-Apr 42 72     |                  | 74 87 75 90       | 35           | 43.4    | 18.6   | 26,778   | 7.02       |
| Iceland          | 3-Apr        | 3-Apr 76 52      |                  | 87 82 90 100      | 1            | 37.3    | 14.4   | 46,483   | 2.91       |
| Ireland          | 26-Apr       | 26-Apr 68 79     |                  | 100 84 93 88      | 9            | 38.7    | 13.9   | 67,335   | 2.96       |
| Israel           | 10-Apr       | 10-Apr 36 52     |                  | 71 76 79 62       | 1            | 30.6    | 11.7   | 33,132   | 2.99       |
| Italy            | 28-Mar       | 28-Mar 76 77     |                  | 70 63 61 66       | 1            | 47.9    | 23.0   | 35,220   | 3.18       |
| Japan            | 23-Mar       | 23-Mar 35 32     |                  | 35 31 30 28       | 1            | 48.2    | 27.0   | 39,002   | 13.05      |
| Korea            | 20-Mar       | 20-Mar 9 9       |                  | 11 10 10 11       | 1            | 43.4    | 13.9   | 35,938   | 12.27      |
| Latvia           | 22-Apr       | 22-Apr 36 51     |                  | 100 80 80 78      | 29           | 43.9    | 19.8   | 25,064   | 5.57       |
| Lithuania        | 4-Apr        | 4-Apr 36 53      |                  | 74 82 95 100      | 35           | 43.5    | 19.0   | 29,524   | 6.56       |
| Luxembourg       | 12-Mar       | 12-Mar 52 62     |                  | 100 97 89 95      | 13           | 39.7    | 14.3   | 94,278   | 4.51       |
| Mexico           | 8-May        | 8-May 10 15      |                  | 38 47 51 54       | 62           | 29.3    | 6.9    | 17,336   | 1.38       |
| Netherlands      | 8-Apr        | 8-Apr 63 65      |                  | 99 91 87 100      | 9            | 43.2    | 18.8   | 48,473   | 3.32       |
| New Zealand      | 14-Apr       | 14-Apr 28 31     |                  | 49 51 59 71       | 43           | 37.9    | 15.3   | 36,086   | 2.61       |
| Norway           | 15-Apr       | 15-Apr 58 80     |                  | 100 73 72 64      | 6            | 39.7    | 16.8   | 64,800   | 3.60       |
| Poland           | 25-Apr       | 25-Apr 61 88     |                  | 100 97 95 82      | 8            | 41.8    | 16.8   | 27,216   | 6.62       |
| Portugal         | 25-Apr       | 25-Apr 73 100    |                  | 100 94 90 79      | 6            | 46.2    | 21.5   | 27,937   | 3.39       |
| Slovak Reb.      | 16-Apr       | 16-Apr 60 65     |                  | 92 88 100 96      | 10           | 41.2    | 15.1   | 30,155   | 5.82       |
| Slovenia         | 6-Apr        | 6-Apr 53 68      |                  | 100 83 76 70      | 7            | 44.5    | 19.1   | 31,401   | 4.50       |
| Spain            | 3-Apr        | 3-Apr 69 75      |                  | 100 98 95 88      | 7            | 45.5    | 19.4   | 34,272   | 2.97       |
| Sweden           | 25-Apr       | 25-Apr 62 83     |                  | 100 78 73 70      | 6            | 41.0    | 20.0   | 46,949   | 2.22       |
| Switzerland      | 1-Apr        | 1-Apr 50 57      |                  | 75 100 79 91      | 14           | 43.1    | 18.4   | 57,410   | 4.53       |
| Turkey           | 20-Apr       | 20-Apr 46 100    |                  | 54 67 51 52       | 20           | 31.6    | 8.2    | 25,129   | 2.81       |
| UK               | 22-Apr       | 22-Apr 51 59     |                  | 90 82 82 87       | 20           | 40.8    | 18.5   | 39,753   | 2.54       |
| United States    | 16-Apr       | 16-Apr 45 57     |                  | 100 90 78 98      | 23           | 38.3    | 15.4   | 54,225   | 2.77       |
was one of the countries that responded most sensitively to the pandemic declaration of the WHO (showing the peak of RSV on March 12), and above all because detailed content analyses of Google searches were available for the United States (Cairo, 2020; Google Trends, 2020), allowing us to understand the changes in the context of searches. Germany was selected from the “insensitive” cluster and France from the “general” cluster; these selections took account of the ease of comparison with the United States based on factors such as economic size, population, and medical infrastructure. Fig. 4 compares the trends in RSV, new cases, and new deaths in the above three countries in the period of 100 days from February 1 to May 11, which is the target period for our analysis. The trends in the United States and Germany form a contrast, with the United States showing a remarkable increase in RSV in the event week, while Germany showed a relatively more sensitive reaction to other factors afterwards. The trends in new cases were also different: Germany’s statistics fell sharply after peaking at the end of March, but the new cases in United States fell slowly. The trends in France were at an intermediate level compared to those in the United States and Germany. There was also a large difference in the trend of new deaths. In early April, France had the largest number, but the death toll fell dramatically soon thereafter. By contrast, the United States had a peak in deaths in late April, and thereafter the death count remained at a relatively steady level. Meanwhile, Germany had relatively fewer new deaths.

4. Research results

4.1. Determinants of RSV

4.1.1. Determinants variables

As explained in Fig. 2, we analyzed significant determinants using the best subset selection method. According to the results presented in Table 3, among the COVID-19 variables, the number of total tests (SQ) was found to be statistically significant in relation to the increase rate of RSV in the event week. Among the control variables, the proportion of the population over 65 years old was the only significant variable. Observing the direction of the coefficient, we found that as both variables decreased, RSV increased significantly. The model presented in Table 3 had an F-value of 14.315 (p-value <0.000), and therefore we judged this to be a statistically significant model. All VIFs (Variance Inflation Factors) were also less than 1.5, which was a strong result. Furthermore, through the distribution of the studentized residual and P-P diagram analysis, we found that the presented model performed pre-processing well for outliers and leverage (James et al., 2013), and therefore we concluded that there was no problem in using the determinants analysis results (model) indicated in Table 2. The explanatory power (adj.R squared) of the model presented in Table 3 was 0.432, demonstrating a high level of explanatory power according to standards of social science research (Cohen, 1988).

For reference, we applied the Lasso method to analyze the same independent variables. Based on these results, three variables (total tests (SQ), aged 65 older, and change of new cases) were judged to significant variables and the coefficient of each were respectively -0.018, -0.003, and 0.025. Compared to the results from best subset selection presented in Table 3, the Lasso results were consistent with the significance and coefficient direction for the two variables that were shown to be significant in Table 3. Although there were differences in the analysis results for the change of new cases, the Lasso method yielded the same results in the case of the second strong model, which we proposed based on the best subset selection method. Based on these findings, we judged that the results from the two methods did not differ widely. Also, as discussed above, we considered the fact that the Lasso method was inadequate for processing the missing values of the variables and that there was an insufficient number of training cases (37 countries) and consequently adopted the best subset selection results shown in Table 3 for our final analysis.

4.1.2. Country-specific differences

According to preceding studies, there are differences by country in the RSV responses to external events related to COVID-19 (Hu et al., 2020). Therefore, there is a limit to generalizing from the hypothesis tests for specific countries. For this reason, to reduce and select the target of our hypothesis tests, we performed clustering analysis prior to analyzing the impacts of the declaration. To cluster the characteristics of countries, we examined the distribution of countries focusing on the total number of tests, which appeared to have a significant effect on changes in RSV, as shown in Fig. 5. In contrast to Iceland, where there was a relatively high number of total tests and almost no change in RSV, most countries were clustered in the third quadrant, centering on Mexico, which is an example of a country where the total test number was relatively small but RSV rose significantly.

To more clearly visualize the differences between countries shown in Fig. 3. Scatter plot of RSV duration and change rate in the event week by country.
Fig. 4. Comparison of trends in RSV, new cases, and new deaths in the USA (top), Germany (middle), and France (bottom).

Table 3
Results of the determinants analysis of the rate of RSV change in the event week

| Model                | Coefficient | Std.  | Importance | t-static | Sig.   | 95% confidence interval |
|----------------------|-------------|-------|------------|----------|--------|-------------------------|
| (Constant)           | 4.415       | 0.481 |            | 9.182    | 0.000  | [5.393, 3.436]          |
| Total tests (SQ)     | -0.495      | 0.141 | 0.530      | -3.498   | 0.001  | [-0.783, -0.207]       |
| Aged 65 older        | -0.089      | 0.027 | 0.470      | -3.297   | 0.002  | [-0.144, -0.034]       |

Fig. 5, Table 4 presents the result of clustering into four clusters according to the DBSCAN method. Table 4 shows that the countries can be divided into 3 clusters excluding outliers. Based on total tests, which is also the most significant determinants variable in relation to our variable of interest (RSV change), the clusters that can be viewed as distinct were as shown in Table 4. Cluster 0 represents the outliers, countries that are not included in any cluster. Most countries belonged to Cluster 1, consisting of countries in which the relationship between the rate of change in RSV and total tests was moderate, which we named the “general cluster.” Cluster 2 is comprised of countries with a relatively high rate of change in RSV which we refer to as the “sensitive cluster,” while Cluster 3 consists of countries with a low rate of change in

K, the number of groups, was set as 4 so that the minimum comparison groups would consist of 3 groups, excluding the outliers.
RSV, which we call the “insensitive cluster.” We selected the United States as the representative country for the sensitive cluster, Germany for the insensitive cluster, and France for the general cluster. Table 4 also provides a country-by-country ranking of the ratio of total tests (SQ) to the RSV change rate in the event week. Germany was ranked 6th within the top group (1-12), France was 7th in the middle group (13-25), and the United States was 5th in the lower ranking group (26-37), and the fact that these three countries were each in the mid-range within each group was also considered in our country selection.

In the selection of representative countries, we also considered the population aged 65 and older, a control variable that showed a significant correlation in Table 3. The variable percentile analysis results for the population aged 65 and older (Hinges of Tukey, excluding outliers) was 15.41% in percentile 25%, 19.68% in 75% and 21.36% in 90%. As seen in Table 2, these values were almost the same as the “aged 65 and older” values of the United States (15.41%), France (19.72%), and Germany (21.45%), respectively.

### 4.2. Effects of RSV

#### 4.2.1. Empirical demonstration of changes in RSV resulting from the pandemic declaration

We performed intervention analysis to statistically examine whether the RSV increase presented in Table 2 was caused by the pandemic declaration, that is, by an external intervention. First, we analyzed the United States, the representative country of the sensitive cluster. Intervention analysis requires the development of a time series model (ARMA) including interventions. Table 5 shows the prediction of the intervention model that would best predict the change in RSV in the United States during the 100 days of the target period of analysis (February 1 – May 11).

---

**Table 4**

Results of clustering according to the DBSCAN method

| Cluster No. | Named cluster | Number of countries | Name of countries |
|-------------|---------------|---------------------|-------------------|
| 0           | Outlier       | 9                   | Australia (9), Canada (26), Chile (36), Iceland (2), Italy (1), Japan (22), Korea (3), Latvia (23), Mexico (37) |
| 1           | General       | 18                  | Austria (18), Belgium (16), Czech Republic (24), Denmark (11), Estonia (17), Finland (21), France (19), Greece (27), Ireland (12), Israel (8), Luxembourg (14), Netherlands (13), New Zealand (28), Portugal (15), Slovak Republic (29), Spain (20), Sweden (10), UK (25) |
| 2           | Sensitive     | 6                   | Colombia (34), Hungary (33), Lithuania (31), Poland (32), Turkey (35), United States (30) |
| 3           | Insensitive   | 4                   | Germany (6), Norway (5), Slovenia (7), Switzerland (4) |

Note: () is the ranking of countries based on the ratio of total tests (SQ) to the weekly change rate of RSV.

---

**Table 5**

Results of the development of the RSV prediction intervention model for the United States

| Variable                      | Coefficient | Std. Error | t-Statistic | Prob. |
|-------------------------------|-------------|------------|-------------|-------|
| X (Intervention effect)       | 20.270      | 1.602      | 12.654      | 0.000 |
| AR(1)                         | 1.146       | 0.074      | 15.414      | 0.000 |
| AR(2)                         | -0.167      | 0.078      | -2.154      | 0.034 |
| MA(3)                         | 0.405       | 0.092      | 4.396       | 0.000 |
| SIGMASQ                       | 22.300      | 2.071      | 10.766      | 0.000 |

Summary Statistics

- R-squared: 0.965
- Mean dependent var: 32.317
- Adjusted R-squared: 0.964
- S.D. dependent var: 25.379
- S.E. of regression: 4.844
- Akaikes info criterion: 6.087
- Schwarz criterion: 6.216
- Log likelihood: -302.374
- Durbin-Watson stat: 1.962

Prob (F-statistic): 0.000

---

8 We did not perform clustering based on the ratio of total tests (SQ) to the rate of change in RSV in the event week because it is difficult to distinguish when the extreme value, the denominator and the numerator are simultaneously small (Japan), as shown in Table 5. We used the DBSCAN method so that we would be able to consider the size as well as the ratio.

9 In Table 5, the variable X which indicates the intervention effect had the value of 1 only on March 12; on other days, it had a value of 0. SIGMASQ is the estimate of the variance of the innovations and Table 5 shows the results calculated by Eviews10.
According to the data on COVID-19 searches in the United States shown in Table 2, the RSV was 100 on March 12, immediately following the declaration on March 11. Compared to the RSV on March 10 and March 11, which were respectively 45 and 57, this represents a very dramatic increase. The effects of these interventions are also confirmed statistically; as shown in Table 5, the time series models explaining the RSV in the United States are the AR(1), AR(2), and MA(3) models and here too, the intervention effect (X) on March 12 was found to be statistically significant.10 According to the intervention model presented in Table 5, the effect of the intervention (coefficient) may be judged to have had the effect of further increasing the RSV by 20.3 on top of the existing trend of increase. This confirms it had the effect of additionally increasing searches by 25.3% on the single day of March 12.

The left side of Fig. 6 shows the intervention model predicted as static and the measured values. It can be seen that the intervention model’s predictive power is strong. On the right, we compare the measured values with the dynamic model in which we assume the effects of intervention did not exist in the intervention model. The RSV of March 12 predicted by the intervention model was close to 80 (left), while the RSV is close to 60 in the non-intervention dynamic intervention model (right), demonstrating again that there is a difference of 20 in the RSV.11 If we observe the cumulative difference between the measured RSV values and the RSV values predicted assuming non-intervention, for the period of 2 months from March 12 to May 11, there is a difference of 21.9%, which again confirms that the effect of the intervention, for the period of 2 months from March 12 to May 11, is significant with a significance level of 10%, forming a clear contrast.

Synthesizing these findings, we can see that the effect of intervention was statistically significant in the United States (sensitive cluster) and France (general cluster), allowing us to accept the first hypothesis, namely that the external intervention by the WHO declaration will strengthen public awareness. As reported in preceding studies, before and after the WHO pandemic declaration, statistics of Google search queries showed that in addition to queries asking “How to get tested for coronavirus?” and after the WHO pandemic declaration, statistics of Google search queries showed that in addition to queries asking “What is a pandemic?” queries regarding “How to get tested for coronavirus?” also increased rapidly (Cairo, 2020). This indicates that the WHO pandemic declaration not only had the effect of enhancing public awareness, leading to simple information searches, but also may have led to actions, as explained in the research model presented in Fig. 1.

According to the data, as shown in Table 7, AR(1) and AR(4) are significant models, and here the intervention effect of March 12 was also statistically significant.13 As indicated by the intervention model presented in Table 7, the effect of intervention (coefficient) additionally increased the RSV by 9.5 above the increase in accordance with the existing trend, and as of March 12, there was an additional rise of 12.6% in the RSV. Although this rise is lower than that of the United States, it confirmed the statistically significant effects of intervention.

In the same manner as Fig. 6, Fig. 7 compares the prediction results of the intervention model and the non-intervention model. When looking at the cumulative difference between the measured RSV value and the RSV predicted based on non-intervention over the two months from March 12, there was a difference of 20.5%, which was similar to that of the United States.

4.2.2. RSV and new tests and cases

4.2.2.1. RSV and new tests. As discussed above, the WHO pandemic declaration had a significant effect on raising RSV, that is, on strengthening public awareness. As reported in preceding studies, before and after the WHO pandemic declaration, statistics of Google search queries showed that in addition to queries asking “What is a pandemic?” queries regarding “How to get tested for coronavirus?” also increased rapidly (Cairo, 2020). This indicates that the WHO pandemic declaration not only had the effect of enhancing public awareness, leading to simple information searches, but also may have led to actions, as explained in the research model presented in Fig. 1.

10 In the residual analysis, because the p-values of the Q statistic were all insignificant at 5%, the null hypothesis that these white noise terms are statistically independent was adopted (failed to reject).

11 Non-intervention model refers to a model in which the effect of intervention is eliminated from the developed intervention model (that is, all x values are zero).

12 Changing the intervention date to March 11 or 13 did not have any statistical significance. The residual analysis in Table 6 also adopted (failed to reject) the null hypothesis that “white noise terms are statistically independent.”

13 In the residual analysis, because the p-values of the Q statistic are all insignificant at 5%, the null hypothesis that these white noise terms are statistically independent was adopted (failed to be rejected).
Table 6
Development results of the RSV prediction intervention model for Germany

| Variable              | Coefficient | Std. Error | t-Statistic | Prob.  |
|-----------------------|-------------|------------|-------------|--------|
| X (Intervention effect)| 2.055       | 3.995      | 0.514       | 0.608  |
| AR(1)                 | 0.962       | 0.013      | 71.408      | 0.000  |
| SIGMASQ               | 80.692      | 3.947      | 20.444      | 0.000  |
| Summary Statistics    |             |            |             |        |
| R-squared             | 0.828       | Mean       | 29.119      |        |
| Adjusted R-squared    | 0.825       | S.D.       | 21.787      |        |
| S.E. of regression    | 9.119       | Akaike info criterion | 7.314 |
| Sum squared resid     | 6149.889    | Schwarz    | 7.391       |        |
| Log likelihood        | –366.341    | Hannan-Quinn criter. | 7.345 |
| F-statistic           | 161.054     | Durbin-Watson stat | 2.249 |
| Prob(F-statistic)     | 0.000       |            |             |        |

Table 7
Development results of the RSV prediction intervention model for France

| Variable              | Coefficient | Std. Error | t-Statistic | Prob.  |
|-----------------------|-------------|------------|-------------|--------|
| X (Intervention effect)| 9.452       | 1.153      | 8.199       | 0.000  |
| C                     | 26.271      | 15.623     | 1.682       | 0.096  |
| AR(1)                 | 1.082       | 0.043      | 24.919      | 0.000  |
| AR(4)                 | –0.129      | 0.041      | –3.106      | 0.003  |
| SIGMASQ               | 28.424      | 2.412      | 11.785      | 0.000  |
| Summary Statistics    |             |            |             |        |
| R-squared             | 0.941       | Mean       | 30.139      |        |
| Adjusted R-squared    | 0.941       | S.D.       | 22.522      |        |
| S.E. of regression    | 5.468       | Akaike info criterion | 6.313 |
| Sum squared resid     | 2870.811    | Schwarz    | 6.443       |        |
| Log likelihood        | –313.829    | Hannan-Quinn criter. | 6.366 |
| F-statistic           | 400.070     | Durbin-Watson stat | 2.158 |
| Prob(F-statistic)     | 0.000       |            |             |        |

Fig. 7. Comparison of intervention model predictions and measured values (left) and comparison with non-Intervention model predictions (right) – France.

Fig. 8 presents comparatively how this interest in testing is reflected in the RSV. In the United States and Germany, searches for COVID-19 overall and specifically for COVID-19 tests showed similar trends, and there was a strong tendency to search for test information more intensively before and after the WHO pandemic declaration (note the sharper peaks). Table 8 shows that there is almost no time lag statistically, and in cross correlation, there is no time lag present and the Granger causality also existed with a very small time lag. This confirms that in the United States and Germany the interest in testing increased almost at the same time, in the same pattern, as the increase in RSV. However, in the case of France, there was a difference in the pattern, as the interest in testing began to appear somewhat later. In the case of France, cross correlation showed the highest correlation with a 6-day lag, and the Granger causality also had the most significant relationship with the same 6-day lag (F-statistic: 6.982, p-value: < 0.000). These results demonstrate that the WHO pandemic declaration caused an increase in COVID-19-related RSV, and while there may be some time lags, specific awareness about COVID-19 testing was indeed included in the pattern.

To determine whether such interest in tests was linked to the action of taking the COVID-19 test, we examined statistics on new tests as shown in Fig. 9. In Fig. 9, the values in the first period indicate the total number of tests up to just before the event week, and the values presented thereafter are the status of weekly new tests. In all three countries, in the event week, more tests were conducted in that one week than the total number of tests preceding. According to the duration values presented in Table 2, which is related to the correlation between RSV and the number of new tests, the period of the steepest increase was one week in Germany, the country which had the shortest duration of 9 days as seen in Table 2. By contrast, France, with a duration of 16 days, showed a two-week spike. The United States, which had the longest duration of 23 days, continued to show a surge for three weeks. This confirms that actual test taking increased rapidly after the rise in interest or perception regarding testing. The form was similar to the continuous response form seen in pulse intervention, where responses manifest continually after the temporary intervention. Based on the above results, we concluded that the increased interest in the tests seen in Fig. 8 was associated with the rapid increase in new tests shown in Fig. 9.

4.2.2.2. RSV and new cases. Observing the relationship between RSV and new cases in the United States shown in Fig. 4, we can see that the change in RSV tends to precede the latter, resulting in some time lag. According to the cross-correlation presented in Table 9, RSV precedes the latter and has a high correlation in the positive direction with a lag of 23 days. The presence of a positive correlation indicates that this correlation may be the effect of rapid testing due to increased awareness. As for Granger causality, Granger causality was observed at the significance level of 0.05 at the 12-day lag, while at the 21-day lag, and a significant Granger causality was observed at the 10% significance level.

We confirmed that the first difference is needed in the unit root test and took the difference and analyzed the VAR model (proceeding to 24 lags, only the values for the interval with the lowest p-value and values for 1 to 3 lag).
Comparison of temporal correlation between COVID-19 and new cases by country.

Table 8
Comparison of the temporal correlation between RSV for COVID-19 overall and RSV for tests, by country.

| Country | U.S. | Germany | France |
|---------|------|---------|--------|
| Cross correlation [lag interval] | 0.962 [0] | 0.948 [0] | 0.829 [6] |
| Granger causality 1 lag F-statistic (Prob.) | 5.197 (0.025) | 18.997 (0.000) | 6.185 (0.015) |
| 2 lag F-statistic (Prob.) | 4.708 (0.011) | 12.197 (0.000) | 2.971 (0.056) |
| 3 lag F-statistic (Prob.) | 3.611 (0.016) | 8.001 (0.000) | 5.852 (0.001) |

Table 9
Comparison of temporal correlation between COVID-19 RSV and new cases by country.

| Country | U.S. | Germany | France |
|---------|------|---------|--------|
| Test | | | |
| Cross correlation [lag interval] | 0.751 [23] | 0.753 [19] | 0.755 [17] |
| Granger causality F-statistic (Prob.) [lag interval] | 1.883 (0.057) | 9.962 (0.000) | 2.321 (0.040) |
| | [11] | [4] | [6] |
| | 1.935 (0.046) | 9.826 (0.000) | 5.215 (0.000) |
| | [12] | [5] | [11] |
| | 1.759 (0.071) | 5.634 (0.000) | 6.936 (0.000) |
| | [13] | [6] | [20] |
| | 1.676 (0.083) | 1.890 (0.045) | 6.220 (0.000) |
| | [21] | [15] | [21] |

5. Discussion

5.1. Discussion of research results

In the cross-sectional study, the variables that were found to strongly influence the rate of change in RSV immediately after the WHO pandemic declaration were total tests and the proportion of the population over 65 based on the best subset model, as shown in Table 3. The direction of influence all appeared to be negative. This can be attributed to the fact that in countries where there were already a high of total tests, the immediate response to the WHO pandemic declaration was weaker because the public awareness was already relatively high. The

Since the case of Germany was rejected in Hypothesis 1, the discovery of new cases due to the increase in RSV should be interpreted as a result of the existing awareness-raising efforts, reflecting the existing trend, rather than as an outcome of the declaration.

17 Since the case of Germany was rejected in Hypothesis 1, the discovery of new cases due to the increase in RSV should be interpreted as a result of the existing awareness-raising efforts, reflecting the existing trend, rather than as an outcome of the declaration.
relatively low rate of increase in RSV when the proportion of the elderly population was high may be attributed to differences in information search activity by age (Perrin, 2015) and may also be explained by the high proportion of the elderly in confirmed cases (Kass et al., 2020).

Therefore, the response of the WHO pandemic declaration is bound to be different from country to country due to differences in total tests and demographic distribution of the population by age group. As shown in Fig. 5, when the total tests are high, the increase rate of RSV is relatively low even in the wake of the pandemic declaration. This difference can be attributed to the fact that where awareness had already been high, and there is no strong surge in interest or intense information searching, as explained in Fig. 1. In Fig. 9, we saw that Germany, which had a relatively high number of total tests, had a smaller increase in RSV when the proportion of the elderly was high may be attributed to differences in information search activity by age (Perrin, 2015) and may also be explained by the high proportion of the elderly in confirmed cases (Kass et al., 2020).

As shown in Table 5, our analysis of the United States, in the cluster of interest, demonstrated that the WHO intervention had a statistically significant effect on the increase in RSV. We found that the RSV immediately increased by 20.3%, heightening public awareness (or attention), and the difference in RSV during the analysis period following the declaration (up to May 11) also indicates that public awareness increased by 21.9%. By contrast, in Germany, which belongs to the insensitive cluster, the declaration of the WHO showed no statistically significant effect on the increase of RSV (refer to Table 6). In France (general cluster), the WHO intervention had a statistically significant effect on the increase of RSV (refer to Table 7), and raised the public awareness which was at 9.5%, a relatively low level; during the analysis period, public awareness also increased by 20.5%. Thus we may conclude that in the cluster of interest and the general cluster, the WHO pandemic declaration increased public awareness by an additional 10-20%, had a positive effect leading to an increase in tests, which in turn enabled the quick identification of new cases.

Regarding new tests, although in Table 8 we were able to statistically explain the increased interest in tests, in Fig. 9, the statistical relationship with RSV was unclear. This phenomenon can be explained by the lack of preparation in terms of test infrastructure. The heightened interest in testing shown in Fig. 8 could not be directly converted into actual testing due to the lack of test infrastructure. As explained in Fig. 9, the increased interest in testing instead led to a slower rise in new tests. Even when searches are made for information regarding COVID-19 or testing, actual access to COVID-19 tests may have varied depending on the country. Fig. 10 presents results which indicates these differences in test accessibility by country, comparing the percentage of positive test results (cases) by country. As shown in Fig. 10, all three countries exhibited a continual rise in the positive rate until the third week after the event week, which confirms the likelihood that all three countries lacked the infrastructure to administer sufficient testing for symptomatic individuals up to the beginning of April.

Fig. 10 can be effectively used to explain the relationship between RSV and new cases, as described in Table 9. As explained in Tables 9 and 10, RSV was found to be significant in explaining new cases in all three countries. Furthermore, we explained that there was a lag of about 3 weeks in the United States and France and a lag of about 1 to 2 weeks in Germany, and these lags shown in Fig. 10 were affected by the shortage of test infrastructure. In Germany, where there was a relatively high number of total tests and where the positive case rate was low, the lag between the trends of RSV and new cases was short.

### 5.2. Policy implications

We can draw several policy implications from these research results. First, while the WHO pandemic declaration also contributed to initiating emergency measures in each country, it also played a critically important role in raising public awareness. In other words, it contributed to public awareness which induced symptomatic patients to take tests, helping to identify confirmed cases early and treat and isolate patients. Therefore, when considering the timing of its pandemic declaration, the WHO should consider not only the trends in new cases and the number of countries with cases (Mackenzie, 2020; Piper, 2020), but also consider its impact on public awareness. In seeking to raise public awareness, the WHO should minimize countries where intervention would be meaningless and maximize the effectiveness of intervention.
by building strategic alliances with each country (e.g., through test infrastructure readiness), and for this purpose, real-time changes in RSV may be useful.

Secondly, this study confirmed that the WHO’s announcement did not have a similar impact on all OECD countries in terms of public awareness, and that the impact was bound to vary depending on the number of tests conducted before the announcement or the number of confirmed cases. Since the WHO cannot make a pandemic declaration at the optimal time for all countries, the efforts of individual countries are also important. Compared to the countries where the increase in RSV was relatively high, those countries where the RSV surge lasted longer will need to make efforts to provide information even more promptly, because the duration in which awareness precipitates information searches was shorter in such countries. As shown in Table 2 or Fig. 3, if the ratio of increase in RSV after the declaration is relatively low, the duration will be short, and therefore public interest will last only a short time. Taking this into account, it would be possible to efficiently control the speed of policy actions and information dissemination.

Thirdly, if the immediate response and duration of information searches indicate the level of awareness that existed in each country up to the immediately preceding time point, RSV can henceforth also be used for post evaluation of each country’s COVID-19 response efforts. For example, Fig. 5 shows that Denmark, Israel, and Japan had below-average numbers of total tests and as seen in Table 2, these countries had a lower RSV increase rate and shorter duration: this may indicate that these countries had already been more successful in raising public awareness.

Finally, because elderly citizens may not engage in searches actively as younger people to improve their awareness, a separate public awareness effort may be needed to reach this demographic group. Table 3 indicated that when the proportion of the elderly population was high, the information search response was low. This may be a country-specific difference or a difference in the information search channels used by specific age groups (Perrin, 2015), but since it may also be a global characteristic of the elderly age group or an indicator of their lack of public awareness, it is necessary to consider policy initiatives to raise public awareness among the elderly and provide appropriate information.

6. Conclusions and discussion of limitations

This study demonstrated that the WHO pandemic declaration increased public awareness and increased searches for information on COVID-19, contributing to the rapid rise of RSV. We also found that this surge in RSV enhanced understanding of COVID-19 and raised interest in testing which in turn may have made a significant contribution to identifying new cases through the COVID-19 tests.

The scholarly significance of this study is that it proposed a method of quantitatively observing the effects of specific external events, more specifically, global policy reactions, on the general public by using the social big data of RSV. It is also significant that this study objectively illuminated the series of processes beginning from awareness and extending to information search and public quarantine activities (testing). In terms of policy and practical application, our findings significantly demonstrated that the WHO’s involvement had a direct impact on the public as well as on policy-making authorities, suggesting the need for reinterpretation of the role of WHO. Another contribution of this study is that it suggests the potential benefits of governments utilizing RSV when implementing policy in response to the outbreak of infectious diseases such as COVID-19.

One limitation of applying the results of this study is that the RSV responses to the WHO pandemic declaration may not only reflect rational judgments, as explained in general human behavior analysis. Risk awareness inevitably includes the component of emotional reactions, and an abnormal rise in information search, referred to as hype (Fenn and Raskino, 2008; Jun, 2012b; Katafuchi et al., 2020), may be reflected in the results. There had already been a demand from the international community for the pandemic announcement since January 2020 (Piper, 2020), which means that the WHO pandemic declaration may have incited fear in a situation where public risk awareness had already been heightened. This hype phenomenon does not significantly change the conclusions of this study, but it will be necessary to develop a new predictive model when making forecasts regarding the period following the subsidence of the emotional disturbance.

Funding source

This research was supported by Korea Institute of Science and Technology Information (KISTI).

Data reference

The raw data of the COVID-19 used in this study was received by OWID (https://github.com/owid/covid-19-data).

CRediT author statement

| Term          | Seung-Pyo Jun | Hyoung Sun Yoo | Jae-Seong Lee |
|---------------|---------------|----------------|---------------|
| Conceptualization | V             | V              | V             |
| Methodology    | V             | V              | V             |
| Software       | V             | V              | V             |
| Validation     | V             | V              | V             |
| Formal analysis| V             | V              | V             |
| Investigation  | V             | V              | V             |
| Resources      | V             | V              | V             |
| Data Curation  | V             | V              | V             |
| Writing - Original Draft | V       | V              | V             |
| Writing - Review & Editing | V       | V              | V             |
| Visualization  | V             | V              | V             |
| Supervision    | V             | V              | V             |
| Project admin  | V             | V              | V             |
| Funding acquis | V             | V              | V             |
| Resources      | V             | V              | V             |

Declaration of Competing Interest

None.

Acknowledgments

Seung-Pyo Jun, the first and corresponding author, acknowledges support from the Creative Challenge Research Program at Korea Institute of Science and Technology Information. The authors would like to express special gratitude to OWID (Our World in Data team) for sharing valuable COVID-19 data for our research and to our reviewers and editors who have generously provided helpful feedback to improve this paper.

References

Althouse, B.M., Ng, Y.Y., Cummings, D.A., 2011. Prediction of dengue incidence using search query surveillance. PLoS Negl. Trop. Dis. 5, e1258.
Althouse, B.M., Weinberger, D.M., Scarponi, S.V., Pitzer, V.E., Ayers, J.W., Wenger, E., Fung, I.C.-H., Dredze, M., Hu, H., 2019. Google searches accurately forecast RSV hospitalizations. bioRxiv, 607119.
Ankerst, M., Breunig, M.M., Kriegel, H.-P., Sander, J., 1999. OPTICS: ordering points to identify the clustering structure. ACM SIGMOD Record 28, 49–60.
Ayyoubzadeh, S.M., Ayyoubzadeh, S.M., Zahedi, H., Ahmadi, M., Kalhori, S.R.N., 2020. Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMR Public Health Surveillance 6, e18826.
BBC news, 2020a. Coronavirus: Germany to Impose Border Controls Over Coronavirus. BBC. https://www.bbc.com/news/world/europe-51897069.
BBC News, 2020b. Coronavirus: Spain and France Announce Sweeping Restrictions. BBC. https://www.bbc.com/news/world/europe-51892477.
Boehm, A., Pizzini, A., Sonnewerth, T., Loefller-Ragg, J., Lamina, C., Weiss, G., Tancevski, I., 2019. Assessing global COPD awareness with Google Trends. Eur. Respir. J. 53.
Thelocal.de, 2020. UPDATE: All German states announce school and Kita closures. Local Germany. https://www.thelocal.de/20200313/coronavirus-bavaria-and-saarland-to-close-all-schools.

Vitzthum, T., 2020. I cannot understand reasons for the exception of Great Britain. Die Welt. https://www.welt.de/politik/deutschland/article206519487/Einreisestopp-fuer-die-USA-Zynischer-Versuch-abzulenken.html.

Walker, A., Hopkins, C., Surda, P., 2020. The use of google trends to investigate the loss of smell related searches during COVID-19 outbreak. International Forum of Allergy & Rhinology. Wiley Online Library, pp. 859-847.

Wang, J., Tang, K., Feng, K., Iu, W., 2020. High temperature and high humidity reduce the transmission of COVID-19. Available at SSRN 3551767.

Wang, Z., Guo, Q., Sun, S., Xia, C., 2019. The impact of awareness diffusion on SIR-like epidemics in multiplex networks. Appl. Math. Comput. 349, 134-147.

WHO, 2020. WHO timeline - COVID-19. World health organization (WHO). https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19.

Woodyard, C., 2020. Q&A: What you need to know about the United Kingdom and Ireland travel ban. USA Today. https://www.usatoday.com/story/travel/airline-news/2020/03/16/coronavirus-uk-ireland-travel-ban-what-know-europe-travel-ban/5060715002/.

Wu, Q., Fu, X., Small, M., Xu, X.-J., 2012. The impact of awareness on epidemic spreading in networks. Chaos 22, 013101.

Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., Du, M., Liu, M., 2020. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Sci. Total Environ., 139051.

Yang, S., Santillana, M., Kou, S.C., 2015. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc. Natl. Acad. Sci. 112, 14473–14478.

Zang, H., 2018. The effects of global awareness on the spreading of epidemics in multiplex networks. Physica A 492, 1495–1506.

Zhang, T., Ramakrishnan, R., Livny, M., 1996. BIRCH: an efficient data clustering method for very large databases. ACM SIGMOD Record 25, 103–114.

Seung-Pyo Jun is a principal researcher of Data Analysis Platform Center at the Korea Institute of Science and Technology Information (KISTI). Also, he is a professor of Science and Technology Management Policy in University of Science & Technology (UST). He received his Ph.D. from the Science and Technology Studies Program at Korea University. His area of interest includes the demand forecasting, emerging technology detecting and new technology adoption analysis. Recently he is focusing on the science technology and SMEs supporting policy using data intelligence.

Hyoung Sun Yoo is a principal researcher at the Korea Institute of Science and Technology Information (KISTI). Also, he is an assistant professor of Science and Technology Management Policy in University of Science & Technology (UST). His area of interest includes demand forecasting, science and technology policy, complex network, and agent-based modeling.

Jae-Seong Lee is a Ph.D. student in Science and Technology Management Policy at University of Science & Technology (UST). He is also a student researcher at the Data Analysis Platform Center in Korea Institute of Science and Technology Information (KISTI). His areas of interest include National Innovation Systems (NIS), R&D project management, R&D collaboration management, innovation policies for SMEs, and Decision-making Support Systems (DSS) based on machine learning and Artificial Intelligence.