Using Internet Search Queries on the Public Health Perception of Risk Towards COVID-19 to Predict Domestic Air-Travel Volume in South Korea

Cheaheon Lim, Yulin Hswen


Abstract

Background
Estimation of future domestic air-travel volume is important for public health officials to identify potential areas of risk of COVID-19. However, information on the demand for travel during the COVID-19 pandemic has not been studied. Internet searches related to the public health perception of risk related to COVID-19 may be an effective method for projecting flight patterns.

Methods
Relative Search Volume data for keyword clusters relating to public health risk perception toward COVID-19 and domestic flight reservations was retrieved from NAVER. We performed a multivariate time series analysis on the RSVs to estimate monthly domestic air-travel volume in South Korea. The time-series lag that yielded the highest correlation coefficient was used to project the South Korean domestic air-travel volume in May 2020.

Results
Our multivariate time series analysis revealed a correlation coefficient of 0.934 (p<0.01), where RSV for public health risk perception had a lag of 0-months and RSV for domestic flight reservations had a lag of 1-month. By this model, we projected travel volume for May 2020 to be 9,463 domestic flights in South Korea in May 2020, and achieved a predictive margin of 1.75%.

Conclusion
NAVER RSV data has the potential to be used as a method for predicting air-travel patterns in South Korea. Internet search queries from NAVER could assist the Korean Centers for Disease Control preemptively respond to expected increases in travel and future cases of COVID-19.

Background

1.1 COVID-19 and Air-travel in South Korea

South Korea has been praised for their efforts to contain COVID-19 as their daily tally of new cases remains lower than 30 cases throughout mid-April and May 2020 (1). The Korean Centers for Disease Control has addressed the risks of the spread of COVID-19 by relying on extensive contact tracing and testing initiatives, as well as imposing a mandatory 14-day quarantine for international travelers (2–4). However, the Korean Centers for Disease Control has not imposed restrictions on domestic air-travel. Globally, more than 93% of the world’s population currently resides in areas with air-travel restrictions due to the risk of transmission of
COVID-19 (5). Additionally, countries like China and those within Europe have placed restrictions on domestic travel (6–8). Yet, as of April 2020, Korean airlines have resumed commercial domestic flights (9). It is unclear if an increase in demand for domestic air-travel will continue in South Korea as the nation enters the Summer Holiday season. For instance, a rise in domestic air-travel has already been seen for April 2020 compared to March 2020 when benchmarked against flight traffic in 2019 (10). Projections for future domestic air-travel volume are needed to better anticipate the potential rise in domestic air traffic and estimate the future risk for COVID-19 outbreaks.

1.2 Public Risk Perception

Public risk perception is a measurement of how much the public characterizes the severity of a risk. The Health Belief Model proposes that incentives to behave in relation to public health are based on the components of perceived susceptibility and severity of a disease (11). Empirical evidence has shown the measure of risk perception to be effective in changing intent and behavior, whereby heightened or lower perceptions of risk are shown to be key determinants of decisions and actions (12). Therefore, a measure of public health risk perception has the potential to identify the public’s behaviors that can affect disease transmission.

Understanding public health risk perception is important for public health officials to gauge future behaviors in order to develop preventative measures that will minimize the impact of infectious diseases such as COVID-19 (13). However, although public health risk perception has been evaluated as an outcome to monitor the effectiveness of public health messaging or cognitive reactions regarding the contagious spread in relation to COVID-19 (14), it has yet to be used as a factor to predict decisions regarding air-travel. Evidence has emerged that as the case trajectory of the COVID-19 pandemic is contained, public health risk perception of the crisis declines (15,16). Therefore, we hypothesize that public health risk perception will have the ability to project future domestic air-travel volume, whereby a decline in perceived risk of COVID-19 will result in a rebound of domestic travel in South Korea.

1.3 Relative Search Volume on NAVER

Previous studies on public health risk perception in South Korea relied on survey methods (13,15,17). Due to resource intensive and long delays in survey reporting of public health risk perception, a novel approach of using Internet search queries to assess public health risk perception has been put forth. Previous studies have used Relative Search Volume (RSV) from Google trends to evaluate public health risk perception for infectious disease outbreaks such as Zika and Ebola (18–20). Gauging public health risk perception using Internet search queries from relative search provides real-time information that is particularly relevant to rapidly evolving infectious disease pandemics such as COVID-19. Most recently, public health risk perception on COVID-19 in South Korea has been assessed using RSV from NAVER, the largest search engine in South Korea that is more popular than Google Korea (14,21). Although this study provided evidence of the ability of RSV on NAVER to capture community health risk perception towards COVID-19 in South Korea, this study did not explore how these patterns of public health risk perception impacts behaviors such as domestic air-travel in South Korea.

As a World Bank study revealed, there has been a significant decline in the demand for travel and tourism from February to May 2020 across the globe (22,23). Although these reductions in demand for travel is likely from widespread global travel restrictions, in the case of South Korea, the lack of restrictions on domestic air-travel by the Korean Centers for Disease Control indicates that governmental measures are less likely to be responsible for the decline in domestic travel. In the present study, we investigated the ability of public health risk perception towards COVID-19 using RSV from NAVER to project future increases in domestic air-travel volume in South Korea. Modeling the potential recovery in domestic air-travel volume in South Korea can better assist the Korean Centers for Disease Control anticipate increases in public mobility and develop effective strategies to prevent future transmissions of COVID-19.

Methods

2.1 Datasets

2.1.1. NAVER RSV Data

We obtained daily Relative Search Volume (RSV) data from NAVER from January 1st 2019 to May 31st 2020 from the NAVER Datalab Application programming interface (API). NAVER is the predominant search engine in South Korea and holds the dominant market share in the Korean search engine landscape (21). We queried the NAVER Datalab API for RSVs of two keyword clusters that were consolidated across all demographic categories. Based on research methods by Husnayain et al. (14) that effectively evaluated public health risk perception toward COVID-19 in South Korea using NAVER, an RSV keyword cluster that included the terms self-quarantine and quarantine; social distancing, distancing, and social distancing was created. We also sought to identify if NAVER RSV for public health risk
perception provided a better real-time estimate for domestic air-travel volume in South Korea compared to NAVER RSV for domestic flight reservations. Therefore, we generated a keyword cluster with terms related to domestic flight reservations to be a proxy for bookings of domestic flights, then compared it to the correlation and lag-time of RSV for public health perception of risk as an indicator for domestic air-travel volume in South Korea. The specific entries in the keyword cluster for domestic flight reservation RSV were selected based on the assumption that the majority of domestic travelers in South Korea search for and purchase airplane tickets through NAVER. This assumption was made because NAVER’s patented flight ticket search program provides consolidated information on available domestic flight reservations when queried for on the main search engine (24). This method was derived from economic studies that have used relative search volume for flights to forecast tourism demand. (25,26). RSV cluster keywords for public health risk perception towards COVID-19 and domestic flight reservations are shown in Table 1.

### Table 1: Keyword clusters queried for on the NAVER Datalab

<table>
<thead>
<tr>
<th>Public health risk perception towards COVID-19 RSV cluster keywords</th>
<th>Domestic flight reservations RSV clusters keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>English Translation</td>
</tr>
<tr>
<td>가가격리</td>
<td>Self-quarantine</td>
</tr>
<tr>
<td>격리</td>
<td>Quarantine</td>
</tr>
<tr>
<td>비스터스</td>
<td>Mask / Face mask</td>
</tr>
<tr>
<td>사회적 거리두기</td>
<td>Social distancing</td>
</tr>
<tr>
<td>사회적 거리두기</td>
<td>Social distancing</td>
</tr>
<tr>
<td>거리두기</td>
<td>Distancing</td>
</tr>
</tbody>
</table>

2.1.2. Domestic air-travel Volume in South Korea

The Korean Airports Corporation reports flight volume data between all major airports on a monthly basis. We retrieved data from the Korean Airports Corporation on the number of flights between Incheon International Airport (ICN), Gimhae International Airport (PUS), Jeju International Airport (CJU), and Gimpo International Airport (GMP) from January 1st 2019 to May 31st 2020.

2.2. Analysis

Descriptive statistics of the Korean Airports Corporation data on monthly flight volume and NAVER RSV data for public health risk perception towards COVID-19 and domestic flight reservations were conducted to understand the temporal trends in relation to COVID-19. A multivariate time-series lag regression was performed to determine the lagged effect (up to 3-months) of the RSV clusters for public health risk perception towards COVID-19 and domestic flight reservations on monthly domestic flight volume in South Korea. The lag period for each RSV data keyword cluster that yielded the highest coefficient of multiple correlation was utilized to project future domestic air-travel volume in South Korea. NAVER RSV data was queried for using Python via the NAVER Datalab API and all statistical analyses were conducted using MATLAB.

Results

3.1. Temporal Trends

Keyword clusters for public health risk perception toward COVID-19 and domestic flight reservations deviated from 2020 to 2019 and are displayed in Figures 1A and 1B. The RSV for public health risk perception towards COVID-19 was at near-zero levels from Jan-June 2019 and peaked thrice overall on January 31st 2020, (37·4 units), February 24th 2020 (100 units), and March 28th 2020 (68·8 units), as shown in Figure 1A. A steady rise in RSV for public health risk perception was seen on Feb 16th 2020 and RSV reached its highest peak on Feb 24th 2020. Following the third peak in late March, the RSV for public health risk perception began a downward trajectory. One exception was on May 28th 2020, when RSV reached a monthly-high of 40·9. However, the trend did not remain and the RSV for public health risk perception decreased to levels lower than 20 and reached 12·4 units on May 29th.

Results for RSV for domestic flight reservations from Jan-June 2019 and 2020 is shown in Figure 1B. RSV in 2019 and peaked on February 25th 2019 at 100 units, whereas on February 25th 2020, the RSV was 28·5 units. In 2020, RSV for domestic flight reservations saw a decline on February 18th, continued to decrease until March 1st and a rise was seen on April 5th with a peak on April 29th. In 2019, RSV for domestic flight reservations was 59·2 units on April 20th, while in 2020 the RSV was 66·3 units. Except for this peak in RSV for domestic flight reservations on April 20th 2020, the RSV for domestic flight reservations remained lower in 2020 compared to 2019.
3.2. Lag-Time Series

Table 2 presents the results of the multivariate lag-time series analysis of the RSV clusters for public health risk perception towards COVID-19 and domestic flight reservations on monthly domestic flight volume in South Korea. The highest coefficient of multiple correlation of 0.934 was observed when lag-month for RSV for keyword cluster on public health risk perception towards COVID-19 was lag=0 (p<.001), while lag-month for RSV for keyword cluster on domestic flight reservations was lag=1 (p<.01).

In comparison to the RSV cluster for domestic flight reservations, the RSV keyword cluster for public health risk perception towards COVID-19 was a more significant predictor for domestic monthly flight volume in South Korea. Each RSV unit increase for the keyword cluster on public health risk perception towards COVID-19 in a given month decreases the number of domestic flights by 45.4 flights (p<.001) in the same month (lag=0), while each RSV unit increase for the keyword cluster on domestic flight reservations increases the number of domestic flights by 6656.3 flights (p<.01) for the processing month (lag=1). Results of the multivariable lag-time series regression is shown in Table 3.
3.3. Projection of domestic air-travel volume in South Korea

The model for the RSV clusters for public health risk perception towards COVID-19 and domestic flight reservations on monthly domestic flight volume in South Korea with the above lag-times is as follows:

Monthly Domestic Air Travel Volume in South Korea = 7,524.6 - 45.4 RSV public health risk perception, lag=0 + 6656.3 (RSV domestic flight reservations, lag=1)

The RSVs used as inputs are as follows: the keyword cluster for public health risk perception towards COVID-19, lag=0 (i.e. May), had a RSV of 22.9 units (p<0.001); and the keyword cluster for domestic flight reservations, lag=1 (i.e. April), had a RSV of 0.447 units (p<0.01).

These analyses were conducted before the Korean Airports Corporation released domestic air-travel data for May 2020. Our model resulted in a projection of 9,463 (SE=668.7) domestic flights in South Korea in May 2020.

Discussion

4.1. Relative Search Volume on NAVER and Domestic Air-travel Volume

Our results showed a parallel (lag=0) relationship between NAVER relative search volume (RSV) for public health risk perception towards COVID-19 and domestic air-travel volume in South Korea. We demonstrate that a higher public health risk perception for COVID-19 leads to lower real-world domestic air-travel among South Koreans. Since the Korean Centers for Disease Control has not restricted domestic air-travel in South Korea, our findings corroborates that a reduction in demand for travel observed throughout
February and March is from increases in the public’s health risk perception towards COVID-19 (15,27). Based on these results and the declining trends of the RSV for keywords on public health risk perception, domestic air-travel will rebound leading up to the Summer months and continue to rise in South Korea for 2020.

Our multivariable lag-time series analysis on RSV also revealed that public health risk perception toward COVID-19 was a more significant predictor of real-world domestic air-travel volume compared to the RSV for domestic flight reservations. Previous studies have documented the accurate forecasting of traveling demands using the RSV for domestic flight reservations. (25,26). However, the onset of COVID-19 around the globe and the first case in South Korea being cited on Jan 20th 2020 (28) may have influenced the predictive ability of RSV for domestic flight reservations to be the most accurate factor for predicting domestic air-travel volume in South Korea. The greater predictive ability and parallel lag-time (lag=0) of the RSV for public health risk perception in projecting domestic air-travel volume in South Korea, compared to the RSV for domestic flight reservations, may be due to the rapidly evolving nature of COVID-19 that leads people to make more last-minute changes to travel plans. For instance, public health risk perception could be associated with flight cancellations and is playing a greater role in determining air-travel volume since many airlines have begun to offer more flexibility for flight changes and cancellations (29). The 1-month lag for RSV for domestic flight reservations may be a reflection of pre-emptive flight bookings a month in advance, but this correlation is outweighed by the public’s perception of risk towards COVID-19, which drives large air-travel cancellations in the same month.

4.2. Future Projections for Domestic Air-Travel in South Korea

This study was conducted before the Korean Airports Corporation published domestic air-travel volume data for May 2020. Our model estimated a volume of 9,463 domestic flights between our selected airports in South Korea during this period. As of July 2020, the Korean Airports Corporation released data and reported that 9,297 domestic flights occurred between our selected airports in May 2020. Our model had a narrow margin of error of 1.75%. This high level of prediction makes NAVER RSV data a promising method for projecting future increases in mobility in South Korea.

4.3. Limitations

Several limitations must be considered. Even though commercial domestic air-travel remains unrestricted in South Korea, airlines have yet to resume flights between Incheon International Airport (ICN) and Gimhae International Airport (PUS) as of May 2020. ICN – PUS is a route that historically attracts hundreds of flights each month, and Busan, the municipality closest to PUS, is a popular tourist location in South Korea for the Summer (10,30). Since our models assume that flights between ICN and PUS will remain at 0 for the coming months, we may be underestimating the recovery in domestic air-travel volume if airlines resume offering flights between the two locations. Second, no adjustments were made to account for social distancing protocols, such as leaving middle seats open, that airlines have been implementing (31). However, this also offers an explanation for the decreased predictive strength of the RSV cluster for domestic flight reservations compared to previous studies that have shown these searches to be highly predictive of future flights but do not take into account the structural changes that airlines have made in light of COVID-19. Moreover, it is important to recognize that our results may not be generalizable for other countries given the unique role that NAVER plays in South Korea for flight bookings.

Conclusions

Our findings demonstrate that Internet search queries on NAVER related to public health risk perception towards COVID-19 offer valuable insight into air-travel mobility patterns in South Korea. Novel variables need to be included that take into account the dynamic nature of the COVID-19 pandemic in order to more accurately predict future population-level behaviors. Changes in the airline landscape of COVID-19 has altered public behaviors and previous models of prediction may not account for these changes. The COVID-19 pandemic has caused rapid variations in public behaviors, making real-time information integral to forecasting real-world outcomes. To estimate future patterns of air-travel mobility, the Korean Centers for Disease Control should integrate the use of NAVER relative search volume for public health risk perception.

References


7. Whitley A, Turner M. These countries have imposed China travel restrictions over the coronavirus. Fortune [Internet]. 2020 Feb 6; Available from: https://fortune.com/2020/02/06/countries-china-travel-restrictions-coronavirus/

8. Kim D. Air carriers resume domestic routes for spring travel season. The Korea Herald. 2020 Apr 16;


27. Kommenda N. How is the coronavirus affecting global air traffic? The Guardian. 2020 Apr 3;


