

RESEARCH ARTICLE

Correlation between National Influenza Surveillance Data and Search Queries from Mobile Devices and Desktops in South Korea

Soo-Yong Shin¹*, Taerim Kim²*, Dong-Woo Seo^{2*}, Chang Hwan Sohn², Sung-Hoon Kim³, Seung Mok Ryoo², Yoon-Seon Lee², Jae Ho Lee^{1,2}, Won Young Kim², Kyoung Soo Lim²

1 Department of Biomedical Informatics, Asan Medical Center, Seoul 138–736, Korea, **2** Department of Emergency Medicine, University of Ulsan, College of Medicine, Asan Medical Center, Seoul 138–736, Korea, **3** Department of Anesthesiology and Pain Medicine, University of Ulsan, College of Medicine, Asan Medical Center, Seoul 138–736, Korea

* These authors contributed equally to this work.

* leiseo@gmail.com



Abstract

Background

Digital surveillance using internet search queries can improve both the sensitivity and timeliness of the detection of a health event, such as an influenza outbreak. While it has recently been estimated that the mobile search volume surpasses the desktop search volume and mobile search patterns differ from desktop search patterns, the previous digital surveillance systems did not distinguish mobile and desktop search queries. The purpose of this study was to compare the performance of mobile and desktop search queries in terms of digital influenza surveillance.

Methods and Results

The study period was from September 6, 2010 through August 30, 2014, which consisted of four epidemiological years. Influenza-like illness (ILI) and virologic surveillance data from the Korea Centers for Disease Control and Prevention were used. A total of 210 combined queries from our previous survey work were used for this study. Mobile and desktop weekly search data were extracted from Naver, which is the largest search engine in Korea. Spearman's correlation analysis was used to examine the correlation of the mobile and desktop data with ILI and virologic data in Korea. We also performed lag correlation analysis. We observed that the influenza surveillance performance of mobile search queries matched or exceeded that of desktop search queries over time. The mean correlation coefficients of mobile search queries and the number of queries with an r-value of ≥ 0.7 equaled or became greater than those of desktop searches over the four epidemiological years. A lag correlation analysis of up to two weeks showed similar trends.

Conclusion

Our study shows that mobile search queries for influenza surveillance have equaled or even become greater than desktop search queries over time. In the future development of

OPEN ACCESS

Citation: Shin S-Y, Kim T, Seo D-W, Sohn CH, Kim S-H, Ryoo SM, et al. (2016) Correlation between National Influenza Surveillance Data and Search Queries from Mobile Devices and Desktops in South Korea. PLoS ONE 11(7): e0158539. doi:10.1371/journal.pone.0158539

Editor: Donald R. Olson, New York City Department of Health and Mental Hygiene, UNITED STATES

Received: November 2, 2015

Accepted: June 17, 2016

Published: July 8, 2016

Copyright: © 2016 Shin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

influenza surveillance using search queries, the recognition of changing trend of mobile search data could be necessary.

Introduction

Syndromic surveillance is defined as a dynamic process of collecting near real-time data on symptom clusters that are suggestive of a biological disease outbreak [1, 2]. With international concerns about emerging infectious diseases, bioterrorism, and pandemics that threaten human or veterinary public health, the importance of syndromic surveillance systems has increased [3–6]. For example, the 2009 H1N1 influenza pandemic highlighted the need for a syndromic surveillance system to inform policy and plan for effective health system responses [4].

Conventional surveillance systems depend on case reporting to report disease activity from sentinel clinics and laboratories. For example, influenza surveillance is recommended to monitor influenza-like illness (ILI) reports and influenza virus infections [5]. Because time delays in case reporting and confirmation can interfere with the early detection of outbreaks or increases in influenza cases in the community [5], digital surveillance could improve both the sensitivity and timeliness of health event detection [5, 7].

Recently, based on the rapid progress of information technology, online resources such as search queries [8–11], Twitter [12–14], Wikipedia access [15, 16], Google AdSense [17], and homepages [18, 19] have been highlighted as promising data sources for influenza surveillance. Of these, internet search queries have been widely used to predict influenza outbreaks. Google Flu Trends, which has recently ended their service, has shown a high correlation with conventional ILI surveillance data [9]. We also developed a surveillance model for Korea using query data from Google [10] and the Korean search engine, Daum [11]. Our previous work showed that the national influenza surveillance data from the Korea Centers for Disease Control and Prevention (KCDC) were highly correlated with search queries in Google and Daum [10, 11]. However, there are some criticisms of digital surveillance, such as Google Flu Trends [20–22]. Digital surveillance cannot be a substitute for conventional disease surveillance systems [10, 20, 22, 23]. In addition, there are the problems of noise consisting of changing user behavior, media-stoked panic, and changes in search engine algorithms [11, 20, 23].

It has recently been estimated that the mobile search volume has surpassed that of desktop searches because of the wide adoption of mobile devices [24–26] and mobile search patterns differ from desktop search patterns [27–32]. However, despite the fact that the volume of mobile searches is rapidly expanding, there is little evidence for the impact of these changes on digital surveillance performance. Previous studies on digital surveillance systems did not distinguish mobile and desktop search queries. The purpose of this study was to compare the performance of mobile and desktop search queries in terms of digital influenza surveillance.

Methods

Study period

The study period ran from September 6, 2010 (week 36) through August 30, 2014 (week 35), which consisted of four epidemiological years (2010/11, 2011/12, 2012/13, 2013/14). Analyses were performed by epidemiological year, defined as the period from week 36 through week 35 of the subsequent year by KCDC [33]. The first day of any epidemiological week is Sunday. Week numbering is sequential, beginning with one, and week one of a Korean epidemiological year is the first week of the year that includes January 1 [33].

Data collection

Mobile and desktop weekly search queries were extracted from Naver Trends, which is freely available [34]. Naver was chosen since it is the largest search engine in Korea, with almost 80% of the Korean search market [35] and other search engines do not distinguish mobile/desktop trends. We downloaded weekly search query data from Naver Trends in comma separated value (CSV) format. This data was assigned a value between 0 and 100 by dividing the number of each combined query by the total number of search queries for a specific period (weekly in this study). Naver did not provide the absolute number of query searches like the other search engine companies. Naver provided this data separately for mobile and desktop, but the exact algorithm was unknown. We provided raw data in the supporting information (S1 File).

KCDC ILI is defined as a fever of 38°C with a cough and/or a sore throat. ILI surveillance consists of 850 sentinel clinics across the nation [33]. These clinics report the weekly percentage of outpatients who meet the case definition of ILI. The virologic surveillance data are weekly laboratory tests showing the positive rates for the influenza virus. This network consists of 91 laboratories across the nation. We downloaded the publicly available data from the KCDC website for the same study period [33]. These data were provided in the “Weekly Sentinel Surveillance Report” document [33], and we manually extracted ILI and virologic data from these documents. In our previous work [11], we performed a survey to gather population search queries related to influenza and we combined the queries from the survey results to reflect people’s search behavior[27]. The resulting queries do not include identifying patient information and are publicly available [11]. A total of 210 combined queries were used for this study [11] (S1 File). All queries with significant values were included in this study regardless of whether they were from desktop or mobile searches.

Statistical analysis

Spearman’s correlation analysis was used to examine the correlation of the mobile and desktop data from Naver Trends with ILI and virologic data from KCDC. We used IBM SPSS Statistics software, version 20 (IBM Corporation, Armonk, NY). Strong correlation was defined as a correlation coefficient r -value of ≥ 0.7 . To assess temporal relationships between Naver Trends and KCDC data for up to two weeks, we performed lag correlation analysis. The creators of disease surveillance models, such as Google Flu Trends, have claimed that the estimations of their models are 1–2 weeks ahead of the reports published by the government [9]. Given the difference in the performance between models for disease surveillance and individual internet time series, we thought that the two-week lag analysis was sufficient. Significance was set at $P < 0.05$.

Results

Correlation between the KCDC ILI data and search queries

Among the 210 combined queries, 14 combined queries had statistically significant correlation coefficients with the KCDC data. The coefficients of correlation of 14 queries between the KCDC ILI data and search queries are shown in Table 1. The correlation coefficients of mobile search queries ranged from 0.390 in the 2010/11 epidemiological year (“Bird flu” in Korean: “조류독감”) to 0.910 in 2013/14 (“Bad cold” in Korean: “독감”) and the correlation coefficients of desktop search queries ranged from 0.291 in 2011/12 (“New flu” abbreviation in Korean: “신플”) to 0.931 in 2011/12 (“Tamiflu” in Korean: “타미플루”). The mean coefficients of desktop search queries were higher than those of mobile search queries in 2010/11. However, since 2012/13, the mean coefficients of mobile search queries have been higher than those of desktops (Table 1, Fig 1). No mobile search queries strongly correlated with the KCDC ILI data in the 2010/11 epidemiological

Table 1. Correlation coefficients between KCDC ILI data and mobile and desktop device search query data from epidemiological years 2010/11 through to 2013/14.

Actual search query	English translation	Mobile search				Desktop search			
		2010/11	2011/12	2012/13	2013/14	2010/11	2011/12	2012/13	2013/14
독감	Bad cold	0.485	0.897	0.778	0.910	0.818	0.863	0.681	0.903
조류독감	Bird flu	0.390	0.504	0.402	0.846	0.903	0.524	0.413	0.842
유행성독감	Epidemiological bad cold	N/A	0.851	0.666	0.805	0.646	0.779	0.518	0.746
플루	Flu	N/A	0.567	N/A	0.700	0.801	0.770	N/A	0.784
H1N1	H1N1†	N/A	0.554	0.571	0.670	0.856	0.630	N/A	0.750
인플루엔자	Influenza	N/A	0.823	0.656	0.880	0.784	0.846	0.731	0.882
Influenza	Influenza (English)†	N/A	0.564	0.638	0.787	0.705	0.838	0.751	0.793
신종독감	New bad cold	N/A	0.784	0.698	0.707	0.382	0.672	0.382	0.510
신종플루	New flu	0.641	0.886	0.788	0.736	0.542	0.802	0.763	0.624
신플	New flu (abbreviation) ‡	N/A	N/A	N/A	0.546	0.574	0.291	N/A	0.326
신종인플루엔자	New influenza	0.443	0.625	0.614	0.664	0.648	0.750	0.538	0.705
돼지독감	Swine flu	N/A	0.539	0.400	0.560	0.757	0.488	N/A	0.628
타미플루	Tamiflu	N/A	0.756	0.846	0.667	0.870	0.931	0.864	0.768
Tamiflu	Tamiflu (English)†	N/A	0.721	0.759	0.734	0.323	0.812	0.740	0.684
Mean coefficient (mean ± SD)		0.490 ± 0.108	0.698 ± 0.144	0.651 ± 0.141	0.729 ± 0.109	0.686 ± 0.179	0.714 ± 0.176	0.638 ± 0.164	0.710 ± 0.153
The number of queries with a strong correlation (r-value ≥ 0.7)		0	7	4	9	8	9	5	9

ILI, influenza-like illness; KCDC, Korea Centers for Disease Control and Prevention; N/A, not applicable due to no Naver data or lack of statistical significance. Naver Trends did not report a value if there were too few searches in a given period; all values are $P < 0.05$ except N/A.

†The query was originally submitted in English. All of the other queries were in Korean.

‡“New flu (abbreviation) (신플)” is the “New flu (신종플루)” abbreviation in Korean.

doi:10.1371/journal.pone.0158539.t001

year. However, in 2013/14, there were 9 strongly correlated queries for both mobile and desktop searches (Table 1).

Correlation between the KCDC virologic data and search queries

The coefficients of correlation between KCDC virologic data and search queries are shown in Table 2. The correlation coefficients of mobile search queries ranged from 0.325 in 2010/11 (“Influenza” in Korean: “인플루엔자”) to 0.861 in 2011/12 (“Tamiflu” in Korean: “타미플루”), whereas those of desktop search queries ranged from 0.280 in 2011/12 (“New flu” abbreviation in Korean: “신플”) to 0.841 in 2013/14 (“Tamiflu” in Korean: “타미플루”). The mean coefficients of desktop search queries were higher than those of mobile search queries in 2010/11. However, since 2011/12, the mean coefficients of mobile search queries have been higher than those of desktops (Table 2, Fig 2). No mobile search queries strongly correlated with KCDC virologic data in the 2010/11 epidemiological year but, since 2011/12, mobile search queries have shown better correlations than those of desktops. The correlation trend of the virologic data was similar to that of the ILI data.

Lag correlation analysis

The lag correlation analysis of up to two weeks showed similar trends to those seen in Tables 1 and 2. The mean coefficients of mobile search queries and the number of queries with a strong correlation became equal or greater than those of desktop searches over time (S1–S4 Figs and S1–S4 Tables).

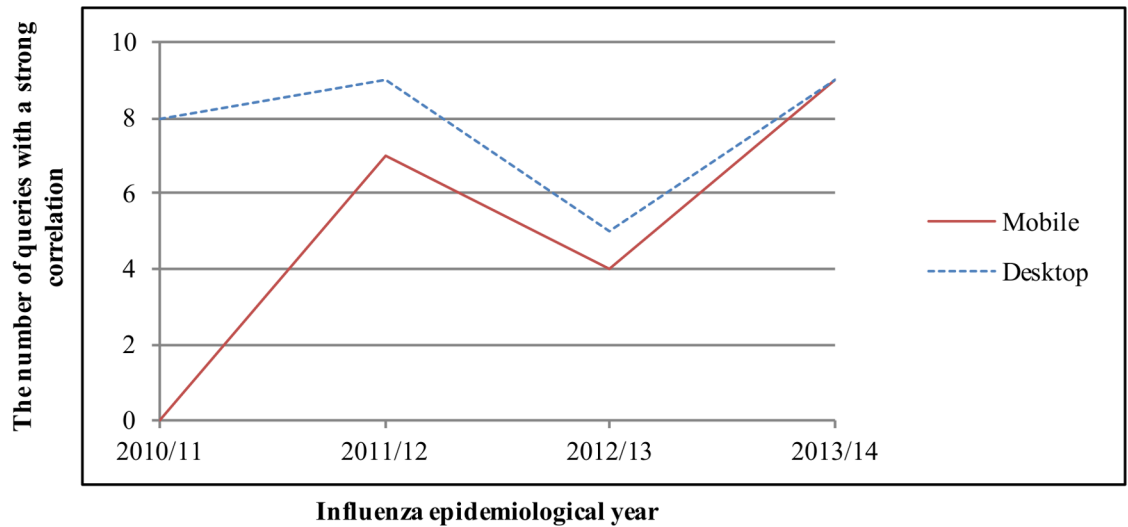


Fig 1. Time series plots of number of search queries having a strong correlation (r -value of ≥ 0.7) with KCDC ILI data.

doi:10.1371/journal.pone.0158539.g001

Discussion

The findings of our present study indicate that the performance of mobile search queries for influenza surveillance has equaled or exceeded that of desktop search queries over time. Many people use internet searches to obtain health information before visiting the hospital [9, 36]. Hence, search query trends can reflect actual disease dissemination earlier than conventional surveillance systems. Previous studies have shown that internet search queries highly correlate with conventional influenza surveillance data [8–11]. However, there are some criticisms about digital surveillance, such as Google Flu Trends [20, 21, 23]. The first is that digital surveillance can be used as complementary source but not as a substitute for conventional disease surveillance systems [10, 20–23]. The second is the noises such as changing user behavior, media stoked panic and search engine algorithm change can affect the performance of digital surveillance systems [11, 22, 23].

In the perspective of changing user behavior, we focused on mobile search queries, because the previous studies did not distinguish mobile and desktop searches [8, 10, 11, 14, 17, 25]. It can be important to distinguish between mobile and desktop search queries for several reasons. First, the mobile search volume is rapidly expanding. In this study, we observed that the performance of mobile search queries for influenza surveillance has improved over time. If this change continues in the future, the importance of mobile searches will increase. Second, the correlation coefficients of all mobile queries except N/A in 2010/11 increased in 2013/14. Because 2010/11 was an early phase of the mobile era in South Korea, there may have been too few searches in 2010/11 (N/A in Tables 1 and 2). It is possible that the increase in mobile search volume was also accompanied by a change in search behavior. Several studies have also reported that mobile and desktop user search patterns differ [27–32].

Noise seems to affect both mobile and desktop searches. In Figs 1 and 2, a decrease in the number of search queries having a strong correlation with KCDC ILI and virologic data for both mobile and desktop searches was observed in 2012/13. We do not know exactly why, but search queries are changing [10, 11, 22, 23]. Moreover, other factors, such as media-stoked panic and changes in search algorithms, may have influenced this [10, 11, 22, 23]. However,

Table 2. Correlation coefficients^a between KCDC virologic data and mobile and desktop device search query data from epidemiological years 2010/11 through to 2013/14.

Actual search query	English translation	Mobile search				Desktop search			
		2010/11	2011/12	2012/13	2013/14	2010/11	2011/12	2012/13	2013/14
독감	Bad cold	0.472	0.722	0.578	0.734	0.553	0.667	0.456	0.626
조류독감	Bird flu	0.601	0.655	0.512	0.736	0.742	0.514	0.516	0.624
유행성독감	Epidemiological bad cold	N/A	0.774	0.632	0.79	0.442	0.673	0.448	0.687
플루	Flu	N/A	0.702	N/A	0.645	0.663	0.805	N/A	0.526
H1N1	H1N1†	N/A	0.486	0.469	0.746	0.710	0.358	0.377	0.633
인플루엔자	Influenza	0.325	0.661	0.722	0.727	0.587	0.629	0.595	0.651
Influenza	Influenza (English)†	N/A	0.584	0.651	0.745	0.687	0.688	0.686	0.747
신종독감	New bad cold	N/A	0.760	0.732	0.717	0.282	0.681	0.395	0.518
신종플루	New flu	0.632	0.782	0.702	0.750	0.421	0.623	0.709	0.650
신플	New flu (abbreviation) ‡	N/A	N/A	N/A	0.559	0.444	0.28	N/A	0.319
신종인플루엔자	New influenza	0.435	0.662	0.677	0.554	0.412	0.541	0.535	0.447
돼지독감	Swine flu	N/A	0.620	N/A	0.672	0.57	0.314	N/A	0.619
타미플루	Tamiflu	0.546	0.861	0.756	0.826	0.719	0.829	0.738	0.841
Tamiflu	Tamiflu (English) †	N/A	0.698	0.731	0.782	0.317	0.797	0.760	0.750
Mean coefficient (mean ± SD)		0.502 ± 0.114	0.690 ± 0.096	0.651 ± 0.095	0.713 ± 0.080	0.539 ± 0.154	0.600 ± 0.178	0.565 ± 0.140	0.617 ± 0.132
The number of queries with a strong correlation (r-value ≥ 0.7)		0	6	5	10	3	3	3	3

ILI, influenza-like illness; KCDC, Korea Centers for Disease Control and Prevention; N/A, not applicable due to no Naver data or lack of statistical significance. Naver Trends did not report a value if there were too few searches in a given period; all values are $P < 0.05$ except N/A.

†The query was originally submitted in English. All of the other queries were in Korean.

‡“New flu (abbreviation) (신플)” is the “New flu (신종플루)” abbreviation in Korean.

doi:10.1371/journal.pone.0158539.t002

despite the probable noise, we observed that the influenza surveillance performance of mobile search queries equaled or exceeded that of desktop search queries over time in this study.

Model output, such as Google Flu Trends, showed high correlation coefficients with conventional surveillance systems for influenza. In Europe, correlation coefficients of 0.716 to 0.940 have been reported for Google Flu Trends [8], and coefficients of 0.80 to 0.99 have been reported in the United States [9]. However, individual internet time series, including this study, cannot be compared directly to model output. In our results with KCDC ILI data, the correlation coefficients of mobile and desktop search queries ranged from 0.390 to 0.910 and from 0.291 to 0.931, respectively. These values are similar to or lower than those reported elsewhere [17, 18]. The different queries may have influenced the performance. Queries used prior to this study only reflected the authors’ opinions [17] or were obtained from databases [9, 18]. To obtain population search queries, we used a survey. In our previous study [11], we performed a survey to gather population search queries related to influenza and combined the queries from the results of the survey to reflect people’s search behavior [27].

There were several limitations to this study. First, the queries were obtained in 2012. Compared to our previous study, fewer queries had statistically significant correlations. This may have influenced the performance. However, we thought that it could be appropriate to use only the 14 statistically significant queries to show the change in the performance of mobile queries for influenza surveillance. Second, queries were not collected by separating the mobile and desktop searches. Several studies have reported that mobile and desktop user search patterns differ [27–32], and a similar result was observed in this study. If specific queries on mobile devices are obtained in additional studies, the surveillance performance may be improved. Third, the first day of a KCDC epidemiological week is Sunday. However, the first day of search

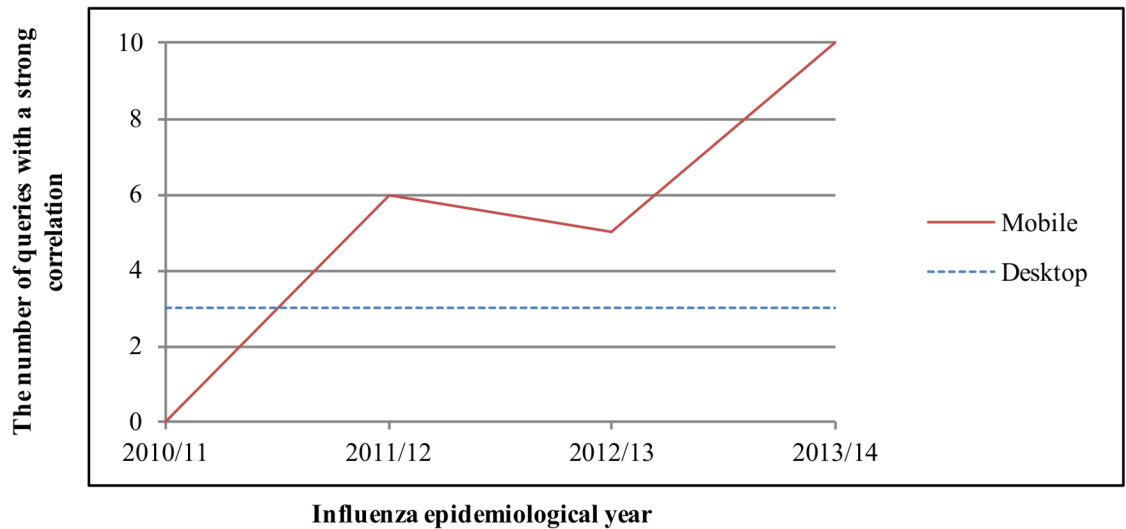


Fig 2. Time series plots of number of search queries having a strong correlation (r-value of ≥ 0.7) with KCDC virologic data.

doi:10.1371/journal.pone.0158539.g002

data from Naver Trends is Monday, and we could not change this. Therefore, it could have skewed the lag analysis. However, we performed the lag analysis using weekly data to minimize the impact. Lastly, we performed this study using open data providing only relative values. Therefore, the results could not be assessed whether the differences were statistically significant, and whether the differences were due to something meaningful or a mechanical artifact.

In summary, we here compared the digital influenza surveillance performances of mobile and desktop search queries. The volume of mobile searches is estimated to surpass desktop searches in the very near future. Our study found that the performance of mobile search queries for influenza surveillance equaled or exceeded that of desktop search queries over time in the study period. In addition, it is possible that the increase in mobile search volume was also accompanied by a change in search behavior. However, we could not show statistically proven difference of correlation or exact causes of these change, since this study was based on limited open data. In the future development of influenza surveillance using search queries, the recognition of changing trend of mobile search data could be necessary.

Supporting Information

S1 Fig. Time series plots of number of search queries having a strong correlation (r-value of ≥ 0.7) with KCDC ILI data from lag correlation analysis (one week preceding of search query).

(TIF)

S2 Fig. Time series plots of number of search queries having a strong correlation (r-value of ≥ 0.7) with KCDC ILI data from lag correlation analysis (two weeks preceding of search query).

(TIF)

S3 Fig. Time series plots of number of search queries having a strong correlation (r-value of ≥ 0.7) with KCDC virologic data from lag correlation analysis (one week preceding of search query).

(TIF)

S4 Fig. Time series plots of number of search queries having a strong correlation (r-value of ≥ 0.7) with KCDC virologic data from lag correlation analysis (two weeks preceding of search query).

(TIF)

S1 File. All combined queries and raw data of this study.

(XLS)

S1 Table. Lag correlation analysis (one week preceding of search query) between search query data and KCDC ILI.

(DOCX)

S2 Table. Lag correlation analysis (two weeks preceding of search query) between search query data and KCDC ILI.

(DOCX)

S3 Table. Lag correlation analysis (one week preceding of search query) between search query data and KCDC virologic data.

(DOCX)

S4 Table. Lag correlation analysis (two weeks preceding of search query) between search query data and KCDC virologic data.

(DOCX)

Author Contributions

Conceived and designed the experiments: DWS. Performed the experiments: SYS TK DWS SMR. Analyzed the data: SYS TK DWS CHS SHK. Contributed reagents/materials/analysis tools: SYS TK DWS. Wrote the paper: SYS TK DWS YSL JHL WYK KSL.

References

1. Triple S Project. Assessment of syndromic surveillance in Europe. *Lancet*. 2011; 378(9806):1833–4. doi: [10.1016/S0140-6736\(11\)60834-9](https://doi.org/10.1016/S0140-6736(11)60834-9) PMID: [22118433](https://pubmed.ncbi.nlm.nih.gov/22118433/).
2. Henning KJ. What is syndromic surveillance? *MMWR Morb Mortal Wkly Rep*. 2004; 53 Suppl:5–11. PMID: [15714620](https://pubmed.ncbi.nlm.nih.gov/15714620/).
3. Irvin CB, Nouhan PP, Rice K. Syndromic analysis of computerized emergency department patients' chief complaints: An opportunity for bioterrorism and influenza surveillance. *Ann Emergency Med*. 2003; 41(4):447–52. doi: [10.1067/mem.2003.104](https://doi.org/10.1067/mem.2003.104)
4. Brownstein JS, Freifeld CC, Chan EH, Keller M, Sonricker AL, Mearu SR, et al. Information technology and global surveillance of cases of 2009 H1N1 influenza. *N Engl J Med*. 2010; 362(18):1731–5. doi: [10.1056/NEJMSr1002707](https://doi.org/10.1056/NEJMSr1002707) PMID: [20445186](https://pubmed.ncbi.nlm.nih.gov/20445186/); PubMed Central PMCID: PMC2922910.
5. Milinovich GJ, Williams GM, Clements ACA, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. *Lancet Infect Dis*. 2014; 14(2):160–8. doi: [10.1016/S1473-3099\(13\)70244-5](https://doi.org/10.1016/S1473-3099(13)70244-5) PMID: [24290841](https://pubmed.ncbi.nlm.nih.gov/24290841/).
6. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature*. 2008; 451(7181):990–3. doi: [10.1038/nature06536](https://doi.org/10.1038/nature06536) 8275740591422483725related:Da0EnThW2XIJ. PMID: [18288193](https://pubmed.ncbi.nlm.nih.gov/18288193/)
7. Morse SS. Public health surveillance and infectious disease detection. *Biosecur Bioterror*. 2012; 10(1):6–16. doi: [10.1089/bsp.2011.0088](https://doi.org/10.1089/bsp.2011.0088) 16112189887464105859related:g2fwK1b_md8J. PMID: [22455675](https://pubmed.ncbi.nlm.nih.gov/22455675/)
8. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks—results for 2009–10. *Euro Surveill*. 2010; 15(29). PMID: [20667303](https://pubmed.ncbi.nlm.nih.gov/20667303/).
9. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009; 457(7232):1012–4. doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634) PMID: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)

10. Cho S, Sohn CH, Jo MW, Shin S-Y, Lee JH, Ryoo SM, et al. Correlation between national influenza surveillance data and google trends in South Korea. *PLoS One*. 2013; 8(12):e81422. doi: [10.1371/journal.pone.0081422](https://doi.org/10.1371/journal.pone.0081422) PMID: [24339927](https://pubmed.ncbi.nlm.nih.gov/24339927/); PubMed Central PMCID: PMC24339927.
11. Seo DW, Jo MW, Sohn CH, Shin S-Y, Lee J, Yu M, et al. Cumulative query method for influenza surveillance using search engine data. *J Med Internet Res*. 2014; 16(12):e289. doi: [10.2196/jmir.3680](https://doi.org/10.2196/jmir.3680) PMID: [25517353](https://pubmed.ncbi.nlm.nih.gov/25517353/)
12. Broniatowski DA, Paul MJ, Dredze M. National and Local Influenza Surveillance through Twitter: An Analysis of the 2012–2013 Influenza Epidemic. *PLoS One*. 2013; 8(12):e83672. doi: [10.1371/journal.pone.0083672](https://doi.org/10.1371/journal.pone.0083672) 038DB0D7-8D56-46D6-B16A-C35B0E46382F. PMID: [24349542](https://pubmed.ncbi.nlm.nih.gov/24349542/)
13. Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, et al. A case study of the New York City 2012–2013 influenza season with daily geocoded twitter data from temporal and spatiotemporal perspectives. *J Med Internet Res*. 2014; 16(10):e236. doi: [10.2196/jmir.3416](https://doi.org/10.2196/jmir.3416) PMID: [25331122](https://pubmed.ncbi.nlm.nih.gov/25331122/)
14. Santos JC, Matos S. Analysing Twitter and web queries for flu trend prediction. *Theor Biol Med Model*. 2014; 11(Suppl 1):S6. doi: [10.1186/1742-4682-11-S1-S6](https://doi.org/10.1186/1742-4682-11-S1-S6) PMID: [25077431](https://pubmed.ncbi.nlm.nih.gov/25077431/)
15. Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global Disease Monitoring and Forecasting with Wikipedia. *PLoS Comp Biol*. 2014; 10(11):e1003892. doi: [10.1371/journal.pcbi.1003892](https://doi.org/10.1371/journal.pcbi.1003892)
16. McIver DJ, Brownstein JS. Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time. *PLoS Comp Biol*. 2014; 10(4):e1003581. doi: [10.1371/journal.pcbi.1003581](https://doi.org/10.1371/journal.pcbi.1003581)
17. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annual Symposium proceedings*. 2006:244–8. PMID: [17238340](https://pubmed.ncbi.nlm.nih.gov/17238340/); PubMed Central PMCID: PMC17238340.
18. Hulth A, Rydevik G. Web query-based surveillance in Sweden during the influenza A(H1N1)2009 pandemic, April 2009 to February 2010. *Euro Surveill*. 2011; 16(18). PMID: [21586265](https://pubmed.ncbi.nlm.nih.gov/21586265/).
19. Hulth A, Rydevik G. GET WELL: an automated surveillance system for gaining new epidemiological knowledge. *BMC Public Health*. 2011; 11:252. doi: [10.1186/1471-2458-11-252](https://doi.org/10.1186/1471-2458-11-252) PMID: [21510860](https://pubmed.ncbi.nlm.nih.gov/21510860/); PubMed Central PMCID: PMC3098167.
20. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Sci*. 2015; 4(1):1–16. doi: [10.1140/epjds/s13688-015-0054-0](https://doi.org/10.1140/epjds/s13688-015-0054-0)
21. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends Data for Detection of Seasonal and Pandemic Influenza: A Comparative Epidemiological Study at Three Geographic Scales. *PLoS Comp Biol*. 2013; 9(10):e1003256–11. doi: [10.1371/journal.pcbi.1003256](https://doi.org/10.1371/journal.pcbi.1003256)
22. Santillana M, Zhang DW, Althouse BM, Ayers JW. What can digital disease detection learn from (an external revision to) google flu trends? *Am J Prev Med*. 2014; 47(3):341–7. doi: [10.1016/j.amepre.2014.05.020](https://doi.org/10.1016/j.amepre.2014.05.020) PMID: [24997572](https://pubmed.ncbi.nlm.nih.gov/24997572/).
23. Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. *Science (New York, NY)*. 2014; 343(6176):1203–5. doi: [10.1126/science.1248506](https://doi.org/10.1126/science.1248506) PMID: [24626916](https://pubmed.ncbi.nlm.nih.gov/24626916/).
24. Mobile now exceeds PC: the biggest shift since the internet began. Available: <http://searchenginewatch.com/sew/opinion/2353616/mobile-now-exceeds-pc-the-biggest-shift-since-the-internet-began>. Accessed May 2, 2015.
25. Mobile search will surpass desktop in 2015. Available: <http://www.emarketer.com/Article/Mobile-Search-Will-Surpass-Desktop-2015/1011657>. Accessed May 2, 2015.
26. 2015 U.S. digital future in focus. Available: <http://www.comscore.com/Insights/Presentations-and-Whitepapers/2015/2015-US-Digital-Future-in-Focus>. Accessed May 2, 2015.
27. Baeza-Yates R, Dupret G, Velasco J. A study of mobile search queries in japan. . . . of the International World Wide Web . . . 2007. doi: [10.1109/IJIS.2008.4497268](https://doi.org/10.1109/IJIS.2008.4497268) 17720386337247934093related: jTZZsnx16_UJ.
28. Church K, Oliver N. Understanding mobile web and mobile search use in today's dynamic mobile landscape. New York, New York, USA: ACM; 2011 Aug 30. 67–76 p.
29. Church K, Smyth B, Bradley K, Cotter P. A large scale study of European mobile search behaviour. *Proceedings of the 10th . . . 2008*.
30. Kamvar M, Baluja S. A large scale study of wireless search behavior: Google mobile search. New York, New York, USA: ACM; 2006 Apr 22. 701–9 p.
31. Li J, Huffman S, Tokuda A. Good abandonment in mobile and PC internet search. . . . conference on Research and development in . . . 2009. 14630338939836676444related:XLmHPMBnCcSJ.

32. Kamvar M, Kellar M, Patel R, Xu Y. Computers and iphones and mobile phones, oh my!: a logs-based comparison of search users on different devices. New York, New York, USA: ACM; 2009 Apr 20. 801–10 p.
33. Korea centers for disease control and prevention. Available: http://www.cdc.go.kr/CDC/info/CdcKrInfo0402.jsp?menuIds=HOME001-MNU1132-MNU1138-MNU0045&fid=84&q_type=&q_value=&pageNum=1. Accessed Nov 2, 2015.
34. Naver trends. Available: <http://trend.naver.com/>. Accessed May 2, 2015.
35. South Korea's internet giant, now or Naver. Available: <http://www.economist.com/news/business/21597937-home-south-koreas-biggest-web-portal-has-thrashed-yahoo-and-kept-google-bay-now-its>. Accessed May 2, 2015.
36. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res*. 2013; 15(7): e147. doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740) PMID: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/); PubMed Central PMCID: PMC3785982.