Original Contribution

Forecasting respiratory infectious outbreaks using ED-based syndromic surveillance for febrile ED visits in a Metropolitan City

Tae Han Kim, MD [a], Ki Jeong Hong, MD, PhD [b,*], Sang Do Shin, MD, PhD [c], Gwan Jin Park, MD [d], Sungwan Kim, PhD [e], Nhayoung Hong [f]

[a] Department of Emergency Medicine, Seoul National University Boramae Medical Center, Republic of Korea
[b] Department of Emergency Medicine, Seoul National University Hospital, Republic of Korea
[c] Department of Emergency Medicine, Seoul National University College of Medicine, Republic of Korea
[d] Department of Emergency Medicine, Chungbuk National University Hospital, Republic of Korea
[e] Institute of Medical and Biological Engineering, Seoul National University, Republic of Korea
[f] Interdisciplinary Program for Bioengineering, Graduate School, Seoul National University, Republic of Korea

ARTICLE INFO

ABSTRACT

Background: Monitoring and detecting sudden outbreaks of respiratory infectious disease is important. Emergency Department (ED)-based syndromic surveillance systems have been introduced for early detection of infectious outbreaks. The aim of this study was to develop and validate a forecasting model of respiratory infectious disease outbreaks based on a nationwide ED syndromic surveillance using daily number of emergency department visits with fever.

Methods: We measured the number of daily ED visits with body temperature ≥ 38.0 °C and daily number of patients diagnosed as respiratory illness by the ICD-10 codes from the National Emergency Department Information System (NEDIS) database of Seoul, Korea. We developed a forecast model according to the Autoregressive Integrated Moving Average (ARIMA) method using the NEDIS data from 2013 to 2014 and validated it using the data from 2015. We defined alarming criteria for extreme numbers of ED febrile visits that exceed the forecasted number. Finally, the predictive performance of the alarm generated by the forecast model was estimated.

Results: From 2013 to 2015, data of 4,080,766 ED visits were collected. 303,469 (7.4%) were ED visits with fever, and 388,943 patients (9.5%) were diagnosed with respiratory infectious disease. The ARIMA (7.0.7) model was the most suitable model for predicting febrile ED visits the next day. The number of patients with respiratory infectious disease spiked concurrently with the alarms generated by the forecast model.

Conclusions: A forecast model using syndromic surveillance based on the number of ED visits was feasible for early detection of ED respiratory infectious disease outbreak.

© 2018 Elsevier Inc. All rights reserved.

1. Background

Infectious disease outbreak is major public health concern and impacts diverse populations. For example, an outbreak of Middle East Respiratory Syndrome (MERS) occurred in Seoul, Korea in 2015 and resulted in mortality and socioeconomic burden [1-4]. Early detection and appropriate public health response is crucial to manage outbreaks and reduce the burden of infectious disease. Several public health surveillance systems have been developed for detection and recognition of infectious outbreaks [5-8]. Monitoring and detecting sudden outbreak is important for outbreak of respiratory infectious disease because of high infectivity, fast propagation, and large burden on health care services.

Emergency Department (ED)-based syndromic surveillance systems have been introduced for faster and earlier detection of several disease categories. Using syndromic parameters derived from clinical data during ED stay instead of confirmed diagnosis or laboratory studies requiring time delay, ED based syndromic surveillance system have shown favorable accuracy and timeliness [7-12].

During the MERS Coronavirus (CoV) outbreak in Korea, the ED was the initial contact point of medical services, and overcrowded EDs were the main place to propagate MERS CoV and delay early detection of outbreak. In Korea, clinical information during ED care is collected into a nationwide electronic database operated by the government called the National Emergency Department Information System (NEDIS). Information collected through the NEDIS network includes vital signs of patients at presentation to ED as well as their final

* Corresponding author at: Department of Emergency Medicine, Seoul National University Hospital, 101, Daehak-Ro, Jongno-Gu, 03080, Seoul, Republic of Korea.
E-mail addresses: emkjhong@gmail.com, ssberg@snu.ac.kr, (K.J. Hong), sdshin@snu.ac.kr, (S.D. Shin), sungwan@snu.ac.kr (S. Kim).

diagnosis [13-15]. However, to detect the outbreak earlier, the development of a forecast model using a nationwide ED database is required.

The goal of this investigation was to develop and validate a forecast model to detect respiratory outbreaks using a real-time nationwide syndromic surveillance system for daily number of febrile ED visits using the NEDIS nationwide ED electronic database. We derived a forecast model for prediction of daily counts of ED visits due to fever and built an alarm system based on differences of predicted and expected values.

2. Methods

2.1. Study setting

This was a retrospective observational study to derive and validate an ED-based forecast model using syndromic surveillance from 2013 to 2015 in Seoul, Korea. Seoul is the capital of Korea, with a population of approximately 10 million and an area of 605.21 km². In Korea, all EDs are designated and classified into 3 levels according to the level and capability of medical service they provide. In Seoul, there was 1 level I ED, approximately 30 level II EDs, and approximately 19 level III EDs included in 2015. The mean number of annual ED visits in Seoul between 2013 and 2015 was approximately 1,200,000. This study was reviewed and approved by an institutional review board (Seoul National University Hospital, IRB No.E-1801-007-911).

2.2. Database

The NEDIS database from 2013 to 2015 was used for the development and validation of an ED-based real-time syndromic surveillance model for febrile ED visits. The NEDIS database system was first established in 2004. All data from patients visiting an ED is collected from the medical records of each institution by a designated and trained coordinator across Korea. Information from each institution is uploaded through the internet to a central server maintained by the National Emergency Medical Center of Korea. Data variables in NEDIS include gender, age, type of insurance, means of transportation, level of consciousness at presentation, emergency operative procedures, time variables (time of ED visit, ED discharge, hospital admission, and hospital discharge), disposition status after the ED encounter, and final ED outcomes (information regarding discharge, transfer, and death). Initial vital signs, including body temperature after ED arrival is measured and collected. Transferring time of each variable in the NEDIS is different according to protocol of the NEDIS quality assessment. Chief complaint and vital signs are transmitted near real-time, while and ED final diagnosis and disposition variables may up to 14 days to transmit.

2.3. Development of ED-based syndromic surveillance for febrile ED visits

To develop a forecast model for respiratory illness outbreak, we used the ED-based syndromic surveillance database during 2013 and 2014. We developed a syndromic surveillance model using the NEDIS database and monitored the daily number of febrile ED visits as a syndromic indicator during the study period. We defined the febrile ED visit by patient records with initial ED presenting body temperature equal or higher than 38.0 °C in the NEDIS, and we collected the final diagnosis coded by the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) of each case defined as a febrile ED visit. If ED final diagnosis of cases corresponded to the ICD-10 codes listed in Table 1, we defined them as patients with respiratory infectious disease. If the daily proportion of patients with designated ICD-10 codes as respiratory infectious disease in syndromic surveillance exceeded the upper 95th percentile of day-of-week mean proportion of respiratory infectious disease calculated during 2013 and 2014, we defined the day as a respiratory infectious disease outbreak.

Table 1
List of ICD-10 codes designated as respiratory infectious disease in the forecast model.

ICD-10 code	Diagnosis
J00	Acute nasopharyngitis [common cold]
J02	Acute pharyngitis (includes sore throat)
J04	Acute laryngitis and tracheitis
J06	Acute upper respiratory infections of multiple and unspecified sites
J09	Avian influenza
J10	Influenza due to other identified influenza virus
J11	Influenza, virus not identified
J16	Pneumonia due to other infectious organisms, not elsewhere classified
J18	Pneumonia, organism unspecified

2.4. Derivation of an ED-based forecast model for respiratory infectious disease outbreak

To develop the forecast model, we conducted a 3-step process. The 1st step was building a forecast model for prediction of the number of febrile ED visits using the number of febrile ED visits of the past few days registered in the syndromic surveillance during 2013 and 2014. The 2nd step was configuring alarm criteria for the extreme number ED febrile visits that exceed the forecasted number using data from 2013 to 2014. The final step was to simulate the forecast model and measure the predictive performance of alarm generated using data from 2015.

To predict the number of febrile ED visits using the forecast model, we used the Autoregressive Integrated Moving Average (ARIMA) modeling method. ARIMA modeling has been used for forecasting time series variables, such as patient volume, in medical science [16,17]. We developed an ARIMA model optimized to forecast febrile ED visits using number of febrile ED visits during the past few days. We developed various ARIMA models and tested the auto-regression, difference, and moving average of each model. Correlograms were plotted and used for comparison between different ARIMA models. AIC and BIC were also measured and compared to select the most adequate ARIMA model with the highest level of fitting.

Alarm was defined according to 2 criteria. If the observed actual number of febrile ED visits exceeded the upper 95th percentile of the expected number of the forecast ARIMA model, we defined the situation as the first alarming criteria. If the observed number of febrile ED visits exceeded the upper 95th percentile of the average number of febrile ED visits according to day-of-week during 2013 and 2014, we defined it as the second criteria. If the daily number of febrile ED visits satisfied both alarm criteria, we activated the forecast alarm. We established 3 alarm durations: 1-day, 3-day, and 7-day alarms. For example, if 7-day alarm is activated, alarming lasts until the 7th day from the day of most recent alarm generation. We simulated and validated the forecast model using ED-based syndromic surveillance from the data of 2015. We also tested the correlation between the number of observed febrile ED visits and patients diagnosed with infectious respiratory disease as their final diagnosis in ED.

2.5. Statistical analysis

Correlation analysis was performed between the daily count of patients with fever and the daily count of patients with ED diagnosis with respiratory infectious disease and plotted simultaneously in a time-series plot. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated with 95% confidence intervals (CI) to measure predictive performance of our forecast model. Separate predictive values were calculated and presented according to length of alarm (1-day alarm, 3-day alarm, and 7-day alarm).

Table 2
Number of ED visits, ED visits with fever, ED visits with final diagnosis of respiratory infection.

Year	Total ED visits	ED visits with fever		ED visits with final diagnosis of respiratory infection	
	N	N	%	N	%
2013	1,238,478	74,003	6.0	109,070	8.8
2014	1,335,698	114,431	8.6	134,059	10.0
2015	1,506,590	115,035	7.6	145,814	9.7
Total	408,0766	303,469	7.4	388,943	9.5

3. Results

From 2013 to 2015, data of 4,080,766 ED visits were collected in the NEDIS (Table 2.). Among them, 303,469 (7.4%) had initial body temperature equal to or higher than 38.0 °C, and 388,943 patients (9.5%) were diagnosed with a designated respiratory infectious disease code (Table 2.) A total of 188,434 ED visits with fever in 2013 and 2014 were used to derive the forecast model. We measured the AIC and BIC values of three different forecasting ARIMA models. The AIC and BIC values of each model were as follows: AIC: 8805.3 and BIC: 8823.7 for ARIMA (1.0.1) model, AIC: 8647.5 and BIC: 8684.2 for ARIMA (3.1.3) model, AIC: 8140.6 and BIC: 8214.1 for ARIMA (7.0.7) model. The ARIMA (7.0.7) model was selected as the final model for predicting ED visits of febrile patients the next day.

The number of daily febrile ED visits and patients with final ED diagnosis of respiratory infectious disease was plotted together in Fig. 2. They showed moderate to strong correlation (Pearson correlation coefficient = 0.86, $p < 0.01$, Fig. 1). The expected number of febrile visits by the forecast ARIMA model was plotted simultaneously with the observed number of febrile ED visits as well as number of patients with infectious respiratory disease diagnosis (Fig. 2).

To validate the forecast model, we simulated our forecast model for ED visit data of 2015. Figs. 3 and 4 show the results of the simulation of the forecast model using ED-based syndromic surveillance throughout the entire year of 2015 and the first quarter of 2015. The number of patients with respiratory infectious disease spiked concurrently with the alarms generated by the forecast model using ED-based syndromic surveillance (Figs. 3, 4).

A total of 29 days of 1-day alarms, 72 days of 3-day alarms, and 133 days of 7-day alarms were generated by the forecast model in 2015 (Table 3). According to length of alarm duration, sensitivity and specificity varied respectively.

4. Discussion

In our study, we developed a forecast model for respiratory infectious disease outbreak using real time ED-based syndromic surveillance monitoring the daily number of febrile patients visiting EDs located in Seoul for 2 years. We retrospectively simulated and validated the derived forecast model using database information from the year 2015. Results of our study suggest that monitoring the number of daily febrile ED visits and alarming the outbreak could successfully predict the surge of patients diagnosed with respiratory infectious diseases in EDs.

A syndromic surveillance system based on the ED database might be beneficial in terms of efficiency and timeliness [7-12]. In many countries, ED data are routinely collected automatically in electronic form with existing ED information systems [18]. Therefore, additional effort for data collection is not needed if a syndromic surveillance system is adopted based on an existing ED health record database. Patients visiting EDs usually have a shorter interval from symptom onset to hospital visit. Abnormal surges in certain disease categories were detected faster when the ED-based syndromic surveillance model was used instead of laboratory or other surveillance models.

In previous surveillance systems, final diagnoses were used as the key element of a surveillance system [7,11,19]. For respiratory infectious diseases, monitoring the number of patients diagnosed with respiratory infectious disease might be a more accurate indicator to predict respiratory infectious disease outbreak than monitoring the number of febrile patients at ED presentation. However final diagnosis at ED is difficult to collect in a timely manner, resulting in time delay of detection for
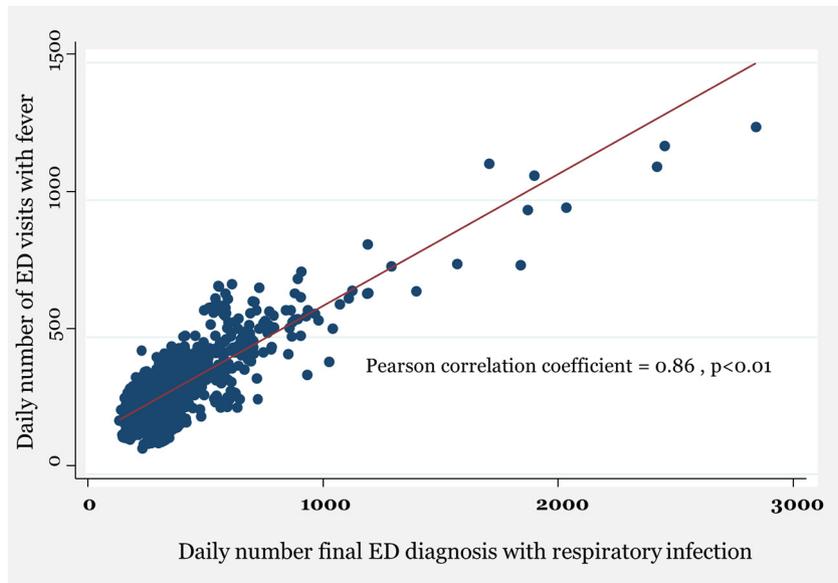

Fig. 1. Correlation between ED visits with fever and ED diagnosis with respiratory infectious disease.
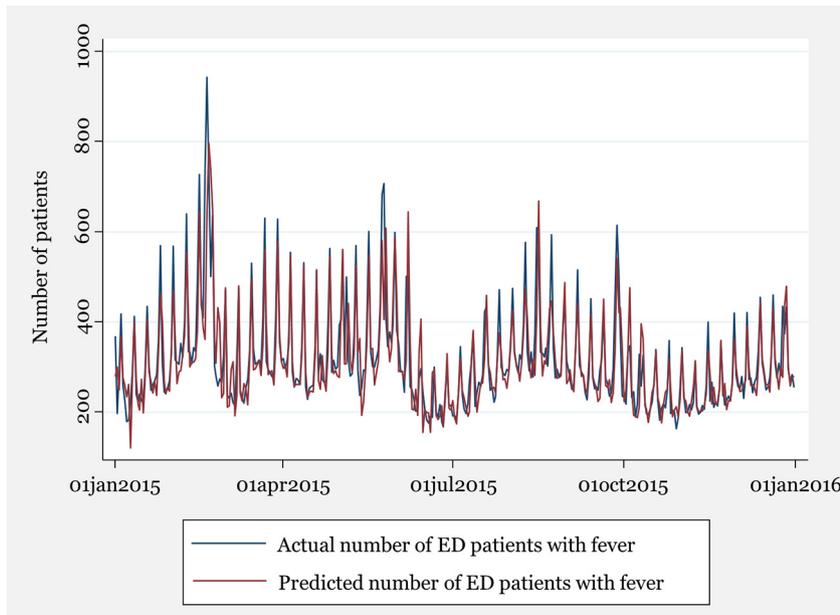
Fig. 2. Observed number and predicted number of daily febrile ED visits from forecast model.

outbreak. In the NEDIS system, time to complete collection of ED final diagnosis information takes up to 14 days from day of ED admission and differs according to each hospital. By contrast, body temperature is checked for all patients during initial triage after ED arrival, and body temperature informati0n is sent to the central NEDIS server directly with triage information without time delay. Body temperature can be used intuitionally for diagnosis or chief complaint without additional data cleaning and coding procedures.

After MERS-CoV outbreak in Korea, thermal scanning devices were installed for the mass screening of febrile patients entering EDs. Until now there has been no universal guideline for screening and alarming based on information obtained from thermal scanning in EDs. The volume of febrile ED patients could be estimated automatically by thermal scanning device, and application of thermal skin device to the forecast

model using ED-based syndromic surveillance may be an efficient method of monitoring the outbreak of respiratory infectious disease.

Although they were not followed by respiratory infectious disease outbreak, a few alarms were generated during early summer and autumn period of 2015 (Fig. 3). We believe that the unexpected extreme number of febrile ED visits was caused by infectious disease categories other than respiratory infectious disease, such as GI tract and urinary tract infection. Further analysis is required to assess inclusions of other infectious diseases with fever that can be detected as a false positive case of respiratory infectious disease in our forecast model.

Various levels of sensitivity and specificity were measured according to different durations of syndromic surveillance alarms (Table 3). Although sensitivity increased according to increase of surveillance alarm duration, specificity decreased with the higher rate of false
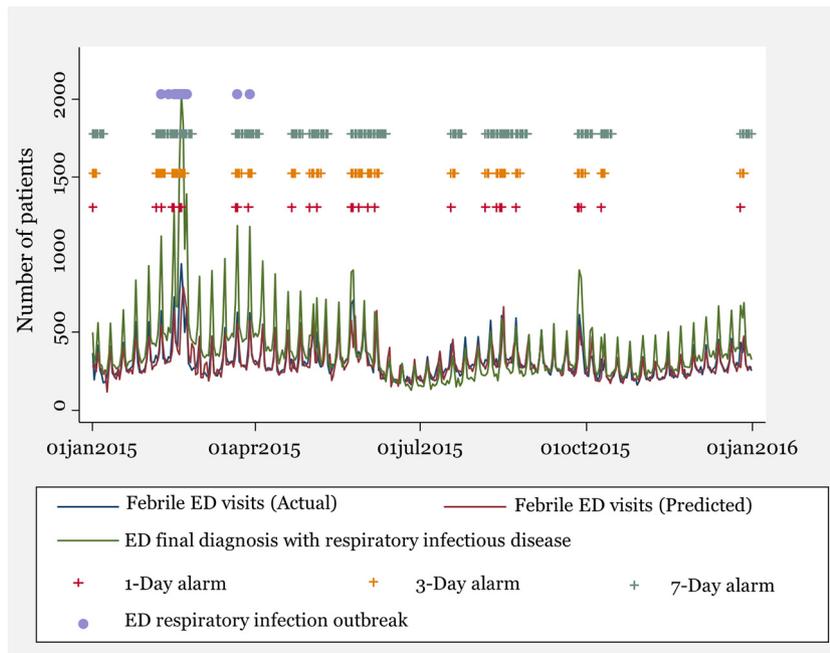

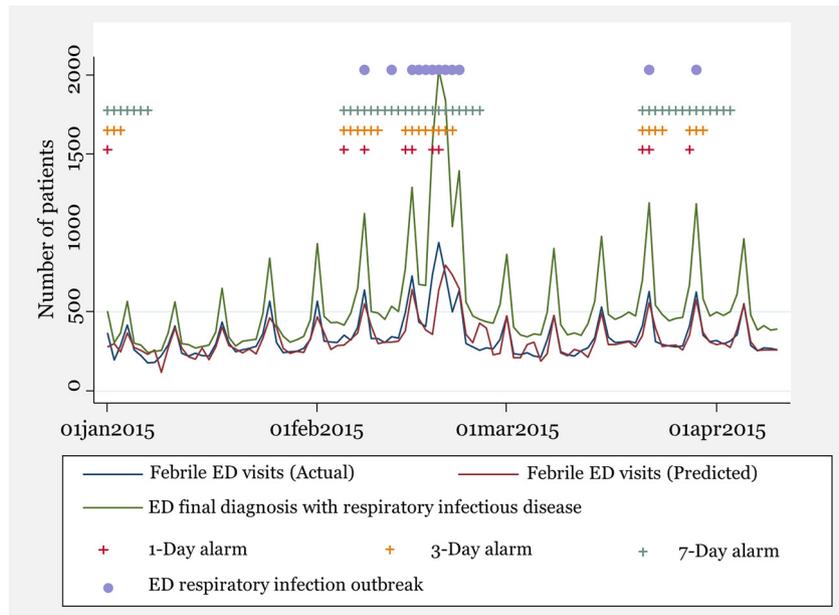**Fig. 3.** Simulation of alarming generated from forecast model to syndromic surveillance during 2015.

Fig. 4. Simulation of alarming generated from forecast model to syndromic surveillance during the first quarter of 2015.

Table 3
Predictive performance of syndromic surveillance model according to length of alarm.

Alarm duration	Total surveillance days	Total days with respiratory infection outbreak		Total days with alarm		Sensitivity	Specificity	PPV	NPV
	N	N	%	N	%	(95% CI)	(95% CI)	(95% CI)	(95% CI)
1-Day alarm	365	12	3.3	29	7.9	0.417 (0.152–0.723)	0.932 (0.901–0.956)	0.172 (0.058–0.358)	0.979 (0.958–0.992)
3-Day alarm	365	12	3.3	72	19.7	0.833 (0.516–0.979)	0.824 (0.781–0.863)	0.139 (0.687–0.241)	0.993 (0.976–0.999)
7-Day alarm	365	12	3.3	133	36.4	1.000 (0.735–1.000)	0.657 (0.605–0.707)	0.090 (0.048–0.152)	1.000 (0.984–1.000)

positive alarms. It is important not to miss the potential time point of respiratory infection outbreak by increasing sensitivity. However, to satisfy the operational purpose of syndromic surveillance, excessive generation of false positive alarms should be decreased to prevent alarm fatigue to laypersons and exhaustion of the public health system. To determine the optimal duration and frequency of surveillance alarms, the performance of predictive power and feasibility to implement into public health system should be considered.

There are a few limitations in our study. First, we could not use the surveillance system for the final diagnosis of respiratory infectious disease based on a culture study report that we could use as a gold standard to assess performance of our forecast model based on syndromic surveillance system. Second, we used 3 years of database values from a Seoul metropolitan city, which results in a limitation of generalizability. Prevalence of respiratory infectious disease and proportion of febrile illness among ED visits could affect the fitness of the ARIMA model or performance of prediction power. Third, we collected diagnosis and vital signs from the NEDIS database retrospectively. Therefore, the quality of the data depends on each institution. However, the National Emergency Medical Center of Korea conducted quality assessment process and certified each hospital for quality of the NEDIS database regularly.

5. Conclusion

We developed a forecast model for the outbreak of respiratory infectious disease using an ED-based syndromic surveillance in a metropolitan city. We monitored the number of febrile ED visits across the country daily based on a real-time syndromic surveillance system using the nationwide ED database network. The expected number of febrile visits by the ARIMA forecast model was identified similarly with observed number of febrile ED visits. The outbreaks of respiratory infectious disease were shown concurrently with the alarms generated by the forecast model using ED-based syndromic surveillance.

References

[1] Paek SH, Kim DK, Lee JH, Kwak YH. The impact of Middle East Respiratory Syndrome outbreak on trends in emergency department utilization patterns. J Korean Med Sci 2017;32(10):1576–80.
[2] Zhang XS, Pebody R, Charlett A, de Angelis D, Birrell P, Kang H, et al. Estimating and modelling the transmissibility of Middle East Respiratory Syndrome CoronaVirus during the 2015 outbreak in the Republic of Korea. Influenza Other Respir Viruses 2017;11(5):434–44.

[3] Chen X, Chughtai AA, Dyda A, MacIntyre CR. Comparative epidemiology of Middle East respiratory syndrome coronavirus (MERS-CoV) in Saudi Arabia and South Korea. Emerg Microbes Infect 2017;6(6):e51.

[4] Ro JS, Lee JS, Kang SC, Jung HM. Worry experienced during the 2015 Middle East Respiratory Syndrome (MERS) pandemic in Korea. PLoS One 2017;12(3):e0173234.

[5] Pfeiffer DU, Minh PQ, Martin V, Epprecht M, Otte MJ. An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data. Vet J 2007;174(2):302–9.

[6] Koopmans M, Wilbrink B, Conyn M, Natrop G, van der Nat H, Vennema H, et al. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. Lancet 2004;363(9409): 587–93.

[7] Beitel AJ, Olson KL, Reis BY, Mandl KD. Use of emergency department chief complaint and diagnostic codes for identifying respiratory illness in a pediatric population. Pediatr Emerg Care 2004;20(6):355–60.

[8] Hughes H, Morbey R, Hughes T, Locker T, Pebody R, Green H, et al. Emergency department syndromic surveillance providing early warning of seasonal respiratory activity in England. Epidemiol Infect 2016;144(5):1052–64.

[9] Hirshon JM. The rationale for developing public health surveillance systems based on emergency department data. Acad Emerg Med Off J Soc Acad Emerg Med 2000; 7(12):1428–32.

[10] Thompson LH, Malik MT, Gumel A, Strome T, Mahmud SM. Emergency department and 'Google flu trends' data as syndromic surveillance indicators for seasonal influenza. Epidemiol Infect 2014;142(11):2397–405.

[11] Travers D, Barnett C, Ising A, Waller A, editors. Timeliness of emergency department diagnoses for syndromic surveillance. AMIA Annual Symposium Proceedings: American Medical Informatics Association; 2006.

[12] Chapman WW, Dowling JN, Wagner MM. Classification of emergency department chief complaints into 7 syndromes: a retrospective analysis of 527,228 patients. Ann Emerg Med 2005;46(5):445–55.

[13] Emergency medical service info-network construction business Available from: http://www.e-gen.or.kr/english/ems_info_network_construction_business.do. [Accessed 03/19/2018].

[14] Kwak YH, Kim DK, Jang HY. Utilization of emergency department by children in Korea. J Korean Med Sci 2012;27(10):1222–8.

[15] Yang HJ, Kim GW, Kim H, Cho JS, Rho TH, Yoon HD, et al. Epidemiology and outcomes in out-of-hospital cardiac arrest: a report from the NEDIS-based cardiac arrest registry in Korea. J Korean Med Sci 2015;30(1):95–103.

[16] Mai Q, Aboagye-Sarfo P, Sanfilippo FM, Preen DB, Fatovich DM. Predicting the number of emergency department presentations in Western Australia: a population-based time series analysis. Emerg Med Australas 2015;27(1):16–21.

[17] Bidargaddi N, Bastiampillai T, Schrader G, Adams R, Piantadosi C, Strobel J, et al. Changes in monthly unemployment rates may predict changes in the number of psychiatric presentations to emergency services in South Australia. BMC Emerg Med 2015;15(1):16.

[18] Farley HL, Baumlin KM, Hamedani AG, Cheung DS, Edwards MR, Fuller DC, et al. Quality and safety implications of emergency department information systems. Ann Emerg Med 2013;62(4):399–407.

[19] Zheng W, Aitken R, Muscatello DJ, Churches T. Potential for early warning of viral influenza activity in the community by monitoring clinical diagnoses of influenza in hospital emergency departments. BMC Public Health 2007;7(1):250.