National Health Alarm Services: Predict and prevent the outbreak of infectious diseases through the combination of various big data

Certificate of Merit with Special Mention, ISSA Good Practice Award – Asia and the Pacific competition 2018

National Health Insurance Service
Republic of Korea

Summary

The National Health Alarm Services (NHAS) of Korea offers various health alarms to the public before the epidemic spreads out rapidly. They combine different data from diverse sources such as national health information from the National Health Insurance Service (NHIS), social networking sites (SNS) such as twitter, blogs, etc., data from Daumsoft Corporation, climate information from the Korea Meteorological Administration (KMA), air pollution information from the Ministry of Environment (MOE), and food poisoning information from the Ministry of Food and Drug Safety (MFDS). Five infectious diseases are continuously monitored, namely, colds, eye diseases, food poisoning, dermatitis and asthma. Data from each organization are merged into one to make this system successful. Data are then analysed for disease prediction using pre-defined formulas, the disease prediction accuracy of which is currently 90 per cent. People can access the alarm services through the Internet or television news easily. This method and system can be adapted by other countries if they have the appropriate data.

The issue or challenge

What was the issue or challenge addressed by your good practice? Please provide a short description.

To prevent unique epidemics in the Republic of Korea using big data is a challenge that confronts the NHIS because epidemics have a huge negative impact on public health as well as on health insurance financing.

The country's four distinct seasons and high population density can lead to some unique epidemics or the spread of infectious diseases in each season. For example, colds and asthma can occur extensively and spread rapidly all over the country during spring and autumn. The rate of food poisoning may increase significantly due to high temperatures in the summer. High levels of ultraviolet rays (UV) in the summer increases the incidence of skin disease. Epidemics such as eye diseases spread rapidly just after school breaks.

Addressing the challenge

What were the main objectives of the plan or strategy to resolve the issue or challenge? List and briefly describe the main elements of the plan or strategy, focusing especially on their innovative feature(s) and expected or intended effects.

The main objectives of developing a disease alarm system were to protect the public from the epidemic of infectious diseases by providing on-time health information using big data.

The innovative elements that make this system a success are as follows:

- active data sourced from various national organizations that are responsible for public interests;

- developing and introducing novel epidemic prediction models based on newly established, innovative algorithms; and

- using strong independent systems to collect various data and simultaneous logic calculations.

We focused on the occurrence cycle of each disease. Major infectious diseases occur at similar intervals within a year. The social phenomena are as follows:

- People visit hospitals more often because of certain diseases.

- People surf the internet with specific keywords for certain diseases.

- The number of patients suddenly increases in a specific geographic area if an epidemic occurs.

- People are searching for weather information more often if the weather is abnormal or forecasted to be bad.

- The number of patients increases when the temperature difference between day and night is large or the season changes.

We can roughly guess what kind of epidemic will occur and when they will occur through these social phenomena. Therefore, information related to the outbreak of a disease such as above mentioned epidemic can be derived if we have correct data. Fortunately, those data have been accumulated for a long time by various organizations such as NHIS, KMA, MOE and MFDS.

We use these strategies to figure out when and where an epidemic occurs.

- Search epidemic related keywords on internet news, blog, and twitter through automatic search bot.

- Refine search results through a knowledge categorizing system to eliminate unnecessary or garbage articles. It increases prediction success rate.

- Combine structured data and unstructured data which are collected from organizations and SNS, respectively.

- Calculate the disease risk index through predefined formula for each infectious disease. If the index level is higher than normal, a disease risk forecast will be made through internet and TV news.

Targets to be achieved

What were the quantitative and/or qualitative targets or key performance indicators that were set for the plan or strategy? Please describe briefly.

The main targets of the NHAS are to predict accurately and to provide the public with forecasts as soon as possible before the spread of an epidemic. By warning in a timely manner, the rapid spread of an epidemic could be prevented.

The other target of this system is to increase the correlation between the prediction model and its actual value. The correlation index is highly related with accuracy.

Regional segmentation and publishing information in a way that is easy for people to understand are also an important goal. The epidemic may occur in a small area and then spread

widely. If indication and warning can be made in small regional units, preparation and treatment will be much easier.

Evaluating the results

Has there been an evaluation of the good practice? Please provide data on the impact and outcomes of the good practice by comparing targets vs actual performance, before-and-after indicators, and/or other types of statistics or measurements.

We have steadily improved and developed our prediction models since 2013. Accordingly, the current correlation indices are over 0.9 (full count is 1), as follows:

- Colds: 0.912
- Eye disease: 0.97
- Food poisoning: 0.965
- Dermatitis: 0.971
- Asthma: 0.933

Now, five major infectious diseases are under control through the alarm and forecast services.

- Alarm: eye disease, food poisoning, and asthma
- Forecast: colds and dermatitis

The forecast of cold and dermatitis can be made up to two days before their occurrence.

The service unit has also been improved. Currently it provides services to 252 local administrative districts. Thus, people can be easily warned by the local administrative units covering areas where they live.

The average number of monthly visitors has continuously increased. It went up from 3,145 persons when the service began, to 265,770 persons per month one year later (Dec. 2016), and 1,725,430 persons two years later (Dec. 2017).

Lessons learned

Based on the organization's experience, name up to three factors which you consider as indispensable to replicate this good practice. Name up to three risks that arose/could arise in implementing this good practice. Please explain these factors and/or risks briefly.

Health insurers in other countries may adapt the Korean model that is based on the concept of analysing data from various data sources to produce new information – the so called "data multiplication".

The NHAS uses climate data, SNS data, weather data and medical treatment data.

- Therefore, there should be adequate data before constructing this kind of prediction system, and the continuous supply of data is essential.

- Weather information based on geographical data is mandatory. Accumulated weather data is advisable. The NHAS uses accumulated data over a number of decades to discern changes in weather patterns.

- Robust algorithms are needed when collecting key words through the internet to remove improper information and refine the results of prediction calculations.

The concept of this system is good and useful. However,

- Broadcasting disease alarms is a relatively sensitive issue. Caution levels should be decided very carefully. The NHAS uses predefined formulas to decide the appropriate caution levels.

- Like disease alarms, disease prediction is also a sensitive matter. It can cause social confusion. Before any formal service is begun, sufficient simulation is required.