Challenges of Trustable AI and Added-Value on Health
B. Séroussi et al. (Eds.)

doi:10.3233/SHTI220533

Influenza Screening Using Patient-Generated Health Data in Post COVID-19 Pandemic

Hyunwoo CHOO [a], Myeongchan KIM [b], Dohyung LEE [b], and Soo-Yong SHIN [a, 1]
[a] *Department of Digital Health, SAIHST, Sungkyunkwan University*
[b] *Mobile Doctor Co. Ltd, South Korea*

Abstract. It is very important to ensure reliable performance of deep learning model for future dataset for healthcare. This is more pronounced in the case of patient generated health data such as patient reported symptoms, which are not collected in a controlled environment. Since there has been a big difference in influenza incidence since the COVID-19 pandemic, we evaluated whether the deep learning model can maintain sufficiently robust performance against these changes. We have collected 226,655 episodes from 110,893 users since June 2020 and tested the influenza screening model, our model showed 87.02% sensitivity and 0.8670 of AUROC. The results of COVID-19 pandemic are comparable to that of before COVID-19 pandemic.

Keywords. Influenza screening, deep learning, mobile health, Patient generated health data

1. Introduction

Reports of seasonal influenza infection worldwide have declined during the COVID-19 pandemic [1]. This does not mean that the influenza virus has disappeared, but the cases of influenza might have been rarely reported influenza-like illness (ILI) surveillance system because of COVID-19 pandemic. In our previous studies, we demonstrated that it was possible to screen influenza using both surveillance information from patient-generated health data (PGHD) and deep learning model [2]. In this study, we will examine whether a deep learning-based screening tool can still operate effectively even in a situation when the pandemic disrupted ILI surveillance system,

2. Methods

We retrospectively collected data through smartphone application Fever Coach from June 2016 to July 2021 in Korea. The variables are as follows: the subject's age, gender, weight, body temperature measured by the user, the type and dose of the drug taken, symptoms, and diagnosis. The proportion of influenza reports for each week was calculated from the collected Fever Coach data (App surveillance). During the same

[1] Corresponding Author, Department of Digital Health, SAIHST, Sungkyunkwan University, 115, Irwon-ro, Gangnam-gu, Seoul, Republic of Korea; E-mail: sy.shin@skku.edu;

period, weekly ILI surveillance provided by the Korea Centers for Disease Control and Prevention (KCDC surveillance) and daily weather data provided by the Korea Meteorological Administration were collected.

We divided the records into episodes and labeled the episodes as influenza or non-influenza depending on whether an influenza report was in the episode. We assigned corresponding KCDC Surveillance, App Surveillance, and the weather data for each episode by the date of the episode. The episodes before June 2020 were used for training, and the episodes after June 2020 were used for test. In the training set, non-influenza episodes were randomly sampled to equal the number of influenza episodes.

An attention-based deep learning model, called multi-time attention network (mTAN), was used for prediction [3]. Adam was used for optimizer, and cross entropy was used for loss function. 5-fold cross validation was applied during the training. All inputs were zero-mean normalized.

3. Results

From June 2016 to July 2021, a total of 149,234 users entered at least one diagnosis record on the Fever Coach app. We collected 226,655 episodes from the data during the period, and divided 110,893 episodes into the training dataset and other 115,762 episodes into the test dataset. Before the COVID-19 pandemic (June 2020), 15.62% of the training dataset were influenza episodes, which decreased 7.16% in the 2019/2020 season, and even dropped to less than 1% in the 2020/2021 season.

Our model for influenza screening achieved a sensitivity 82.4 % (±0.16), AUROC 0.898 (±0.004) in the training set. In the test set, the model achieved a sensitivity 87.02% (±0.15), and AUROC 0.8670 (±0.002).

4. Discussion and Conclusion

In a previous study, we showed the possibility of combining deep learning-based influenza screening using patient reported symptoms. There were influenza outbreaks in all three years we previously analyzed. Therefore, the model was neither trained nor tested for years where there was no influenza outbreak. In this study, we found the model maintained in fair performance even though little influenza case has been reported due to the unprecedented COVID-19 pandemic. Further, we showed that deep learning model using PGHD, which are not well-controlled, can have strong performance.

References

[1] Olsen S, Azziz‐Baumgartner E, Budd A, Brammer L, Sullivan S, Pineda R et al. Decreased influenza activity during the COVID‐19 pandemic−United States, Australia, Chile, and South Africa, 2020. American Journal of Transplantation. 2020;20(12):3681-3685.

[2] Choo H, Kim M, Choi J, Shin J, Shin S. Influenza Screening via Deep Learning Using a Combination of Epidemiological and Patient-Generated Health Data: Development and Validation Study. Journal of Medical Internet Research. 2020;22(10):e21369.

[3] Shukla S, Marlin B. Multi-Time Attention Networks for Irregularly Sampled Time Series [Internet]. arXiv.org. 2022 [cited 23 January 2022]. Available from: https://arxiv.org/abs/2101.10318