Case Study #6 – Vaping on FHIR

CS6440 Intro Health Informatics, 2020 Spring

Xia, Hui

903459648

Georgia Institute of Technology

1. **How could FHIR be used to help identify patients with Vaping-Associated Lung Disease?**

   Vaping-Associated Lung Disease is just one of many atypical, or rare, diseases. Since the introduction of the electronic health record (EHR), patient symptom analysis become more relied on electronic data. That is, "in the current digital age, the electronic health record represents a massive repository of electronic data points representing a diverse array of clinical information" (Carter, 2008). Initiatively, data from EHRs that are shared between many healthcare providers can surpass any medical doctor's personal experience. For the case of rare diseases, the ability of using EHR data to support diagnosis become even more important. This is because, the process of diagnosing the first case of a rare disease usually takes extended time, extra cost of finical and man power, or even life-cost tragedies (Therapies, 2013). Thus, computational technology that could utilize these precious experience well to help recognize rare disease become important in saving time, cost, and lives. Thanks to the recent development of computational technologies, several methods could be used to identify patients with Vaping-Associated Lung Disease:

   A. **Machine learning and big data on FHIR records could also be used to identify patients with Vaping-Associated Lung Disease, or other rare diseases in general.** As rare diseases being rare, the healthcare providers tend to have limited experience to correctly diagnosing them. That is, these is a high possibility that the healthcare providers make faulty judgment, and possibility treat the patient in non-optimized manner. For example, it is usually hard to distinguish SARS from regular pneumonia, without properly built protocols (Jernigan, Low, & Helfand, 2004). To ensure fast and accurate diagnosis of rare diseases, FHIR data on symptoms of rare diseases should be stored in a database, and studied by a machine learning algorithm. Based on the machine learning algorithm, patterns on whether any given symptoms are associated with rare diseases should be brought up for notice. However, this is not easy task, as the size of health care data tend to be large to collect enough rare disease cases. Thus, big data technology could be used to recognize patterns from such large database (Chawla & Davis, 2013).

   B. **Machine learning and big data on FHIR records could also be used to identify potential patients with high risk of Vaping-Associated Lung Disease.** Being an unhealthy life habit, vaping could slowly deteriorate individual's health, without directly causing Vaping-Associated Lung Disease. However, these individuals are exposed to high risk of Vaping-Associated Lung Disease, and their healthcare providers, as well as the patient him/herself, should be notified for such risk. Studies have revealed that vaping individuals will show certain symptoms, such as throat hit, without being diseased (Javed et al., 2017; Li, Zhan, Wang, Leischow, & Zeng, 2016). With FHIR documenting such symptoms, and proper set up of big data analysis, a machine learning algorithm can identify potential patients with Vaping-Associated Lung Disease on their physician visits, and inform such

information to the corresponding healthcare provider. Doing so could be helpful in educating the patient to quit vaping, as patients tend to listen and follow the suggestions made by the PCP and nurse they are familiar with, rather than following other information source (Haas, Leiser, Magill, & Sanyer, 2005).

2. **In addition to FHIR, what technologies would be needed to analyze medical charts to identify these patients automatically?**

As the risks correlated to vaping are often hidden, and subtle, and appears in a time-extended manner (i.e. longer vaping history may result in higher health risk), big data technology and applications could be used to inform the providers about the emerging risks. Recently, the researchers and healthcare providers have been discussing the possibility applying big data I analyzing medical charts generated from patient's visits to clinics, and help to make diagnosis and treatments decisions (Hernán & Robins, 2016; Raghupathi & Raghupathi, 2014). The essentials of applying big data in clinical trials is pattern recognizing. To determine that if any of the medications prescribed to the patients could be associated with health risks, we need to recognize patterns from the patent's medication and health history. That is, data from all medical charts of the patient should be stored in a database. Based on the database, patterns on whether any given prescriptions are associated with health risks should be brought up for notice. However, as we have discussed above for **Question 1**, this is not easy task, as the size of health care data tend to be large, and the number of patients is also large. Big data equipped with proper machine learning algorithm should be investigated to recognize patterns from such large database.

3. **Provide us with examples or ideas you have for communicating the risks of vaping to young people most effectively?**

   A. **Use natural language processing algorithms to analyze social media information**. Although that Medical studies have discovered many e-cigarette may cause adverse health events, currently, most existing clinical trials on the health risk of e-cigarettes are with limited results due to their small sample size and short duration (Callahan-Lyon, 2014; Chen, 2013). On the other hand, the popularity of social media has increases, and many e-cigarette consumers are exchanging their information and experience on social media. Thus, analyzing social media information on vaping become a valuable source for understanding e-cigarette user behavior, health effects, and risks. For example, Xie et. al. implemented a Bidirectional Long Short-Term Memory Recurrent Neural Network on discussions between e-cigarette users, and identified 1591 unique adverse events from their research (Xie, Liu, & Dajun Zeng, 2018). Such information could be promoted when communicating the risks of vaping to young people.

   B. **Use big data to educate the young people on risking and alternatives of vaping.** Currently, big data have been applied on personalized internet search engine (e.g. for advertisement purpose) (Couldry & Turow, 2014). Thus, it is rather feasible to "catch and

educate" young people that has a history of vaping (such data can be acquired using the natural language analysis model described above in **Section 3A**), then show them alternatives of vaping and/or suggestions on how should they quit vaping, when they are looking for e-cigarette products online on internet-based searching engine such as Google.

References

Callahan-Lyon, P. (2014). Electronic cigarettes: human health effects. *Tobacco control, 23*(suppl 2), ii36-ii40.

Carter, J. H. (2008). *Electronic health records: a guide for clinicians and administrators*: ACP Press.

Chawla, N. V., & Davis, D. A. (2013). Bringing big data to personalized healthcare: a patient-centered framework. *Journal of general internal medicine, 28*(3), 660-665.

Chen, I.-L. (2013). FDA summary of adverse events on electronic cigarettes. *Nicotine Tob Res, 15*(2), 615-616.

Couldry, N., & Turow, J. (2014). Advertising, big data and the clearance of the public realm: marketers' new approaches to the content subsidy. *International Journal of Communication, 8*, 1710-1726.

Haas, L. J., Leiser, J., Magill, M. K., & Sanyer, O. N. (2005). Management of the difficult patient. *American family physician, 72*(10), 2063-2068.

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology, 183*(8), 758-764.

Javed, F., Abduljabbar, T., Vohra, F., Malmstrom, H., Rahman, I., & Romanos, G. E. (2017). Comparison of periodontal parameters and self-perceived oral symptoms among cigarette smokers, individuals vaping electronic cigarettes, and never-smokers. *Journal of periodontology, 88*(10), 1059-1065.

Jernigan, J. A., Low, D. E., & Helfand, R. F. (2004). Combining clinical and epidemiologic features for early recognition of SARS. *Emerging infectious diseases, 10*(2), 327.

Li, Q., Zhan, Y., Wang, L., Leischow, S. J., & Zeng, D. D. (2016). Analysis of symptoms and their potential associations with e-liquids' components: a social media study. *BMC Public Health, 16*(1), 674.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science systems, 2*(1), 3.

Therapies, S. H. G. (2013). Rare Disease Impact Report: Insights from patients and the medical community. *Shire Human Genetic Therapies, Tech. Rep.*

Xie, J., Liu, X., & Dajun Zeng, D. (2018). Mining e-cigarette adverse events in social media using Bi-LSTM recurrent neural network with word embedding representation. *Journal of the American Medical Informatics Association, 25*(1), 72-80.