

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

- 1.1. Computer technologies such as machine learning can be used to recognize symptoms of the patients. Using a FHIR database built based on analysis of symptom of Vaping-Associated Lung Disease patients, and under the help of proper specialist whom could provide quantitative labeling of respective symptoms, a supervised learning classifier-based algorithm could be built. Such algorithm could help to screen patients with the possibility of suffering from Vaping-Associated Lung Disease. Various studies have used machine learning tools to help doctors to screen such disease, or other rare diseases. For example, computed tomographic scan images of the chest from patients with Vaping-Associated Lung Disease has been collected. Specific image patterns related with Vaping-Associated Lung Disease, has been discovered as “basilar-predominant consolidation and ground-glass opacity, often with areas of lobular or subpleural sparing.” (Henry, Kanne, & Kligerman, 2019) After more consensus on image patterns are made, and enough data are collected, a supervised machine learning classifier could be built to perform automatic evaluation of the potential risk of Vaping-Associated Lung Disease.
- 1.2. Beyond using diagnosis algorithms to analysis measurable symptoms from FHIR information, patients with Vaping-Associated Lung Disease can also be identified using data from social media. Xie et. al. implemented a Bidirectional Long Short-Term Memory Recurrent Neural Network, which could collect on-line discussions from e-cigarette users (Xie, Liu, & Dajun Zeng, 2018). If vaping experience can be found in a patient's social media history, the patient's doctors should take the risk of Vaping-Associated Lung Disease into consideration.

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

- 2.1. Based on the potential patient's EHR information, the health provider could use big data to estimate a given patient's risk on Vaping-Associated Lung Disease. It has been suggested by studies that vaping will result in certain symptoms, such as throat hit. On the other hand, it has been found that people of healthy lifestyle tend to be less affected by sore throats (Javed et al., 2017; Li, Zhan, Wang, Leischow, & Zeng, 2016). Thus, if the FHIR charts track on the throat condition history for an individual, it is possible to use a computational system

updated for diagnostics of Vaping-Associated Lung Disease. For example, Anbarzadeh and Davari have implemented a fuzzy algorithms to help identify different sore throat sensitivities and provide advises on treatment methods accordingly (Anbarzadeh & Davari, 2015). Such algorithms can be used to identify potential Vaping-Associated Lung Disease patients from the ones that are affected by throat hit or sour throat.

- 2.2. Healthcare providers could also combine the potential patient's search engine history and big data analysis to automatically identify potential patients. To date, online search engines have widely adopted big data and machine learning methods (e.g., advertising campaigns) (Couldry & Turow, 2014). Therefore, it is possible to use find patient's personal on-line search habit, and add these data into analysis. For example, upon the potential patient searching for "vaping" or "e-cigarette" on internet, the searching engine help to feed such information to health providers, who can take the risk of Vaping-Associated Lung Disease into consideration when evaluating this individual's throat complications.

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

- 3.1. Similar to what I have discussed in Section 2.2, healthcare providers could combine the potential (young) patient's online search behavior with big data. Based on online search behavior, it is possible to filter young people that has a history of vaping, and suggest them alternatives of vaping, or simply to quit vaping, when they are looking for purchasing e-cigarette products, or the address of a vaping shop online.
- 3.2. Furthermore, big data can be combined with natural language analysis algorithms to find the potential risks of vaping. Recently, on line information about vaping and e-cigarette consumption has increased. Many e-cigarette consumers exchange their opinions and experience about vaping on internet forums. Thus, big data analysis on social media about vaping provide necessary insights on how vaping affects the population's health. The Bidirectional Long Short-Term Memory Recurrent Neural Network that I have mentioned in Section 1.2 can be used to analysis the discussions between e-cigarette users, and identified 1591 unique adverse events from vaping (Xie et al., 2018). These information could deliver information on risk of vaping effectively to young people.

References

- Anbarzadeh, S., & Davari, H. (2015). Web based fuzzy diagnosis system and evaluation for five diseases with sort throat symptom. *Ciência e Natura*, 37, 239-246.
- 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.
- Henry, T. S., Kanne, J. P., & Kligerman, S. J. J. N. E. J. o. M. (2019). Imaging of vaping-associated lung disease. *381*(15), 1486-1487.
- 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.
- 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.
- 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.