International Journal of Population Data Science

Journal Website: www.ijpds.org

Observational Data Exploration Via Online Tool For For Drugs and Cancer Risk

Iqbal, Usman^{1*}, Nguyen, Phung Anh¹, Syed-Abdul, Shabbir¹, Jian, Wen-Shan¹, and Li, Yu-Chuan Jack¹

¹Taipei Medical University

Objective

Rapid change in health information technology system had dramatically increased health data accumulated. We aimed to develop an online informatics tool in order to evaluate the risk of drugs for cancer by utilizing medical big data.

Approach

We use the Taiwan's National Health Insurance Database that has provided a huge data which covered all health information including characteristics and all drug information i.e. prescriptions, etc. of 23 million Taiwanese population. Front-end development: Web-based interface was developed by using PHP package and Javascript. In addition, we included the guidelines of evidence based medicine (EBM) level 3 for observational study such as cohort, case-control, and/or case serial self-control in order to support users interact with system. Back-end development: A package of Apache, MySQL & PHP was used to build the serve-side of the system. We integrated the Elasticsearch API5 to our system in order to search and analyze data immediately. The example of data transform to person-level from Taiwan NHI database is shown in Box 1. After then, we also integrated the analytics package (ie. R package) to perform the statistical analysis to a given study.

This online analytical tool has capability to massively explore and visualize big data for long term use drugs and cancers through OMOSC system which will help to do mass online studies for long term use drugs and cancer risk. It would help to direct the health care professionals with lack of datamining skills to lead the study. The constructed online system would generate automatically case and controls by utilizing large databases for long term drug exposures and cancer risk.

Results

The results are shown in odds ratio (OR) and if selected some confounding factors then could also get adjusted odds ratio (AOR) for risk estimation with 95% Confidence Intervals (CI). We used SAS statistical software on the same dataset to validate the OMOSC system results. It could help to do massive online studies which will saves time and cost effective.

Conclusion

Since the clinical trials are impossible to conduct due to cultural, cost, ethical, political or social obstacles. Therefore, this kind of research model would play an important role in health care industry by providing an excellent opportunities for solving the technological, informatics, and organizational issues towards other broad domains of drugs evaluation by utilizing large-scale databases.

*Corresponding Author:

Email Address: usman.iqbal85@gmail.com (U. Iqbal)