Call for Contributions:

A Special Issue of Journal of Data Science on

Data Science Approaches to Vaccine Effect Analysis

Deadline: June 30, 2022

Journal of Data Science (`http://jds-online.org`) invites submissions to a special issue on "Data Science Approaches to Vaccine Effect Analysis."

Vaccine is an essential tool for stopping the spread of COVID-19. Evaluation of the effectiveness of the vaccines is important for the success of vaccines. Vaccine effect analysis consists of assessing the efficacy of vaccines in clinical trials and the effectiveness of vaccines under real world conditions. The Centers for Disease Control and Prevention (CDC)'s guidance on vaccine effectiveness analysis include (1) analysis for specific subpopulations, (2) reducing the risk of infection, (3) protection against mild COVID-19 illness, (4) preventing hospitalization, (5) reducing the spread of illness, (6) assessing duration of protection, (7) evaluating the impact of the virus new variants, (8) evaluating the effectiveness of a single dose and delayed second dose, and (9) evaluating impact of population host factors.

The effect analysis may include two kinds of study designs, (1) randomized clinical trials and (2) observational studies. Since the infection and progression of COVID-19 are stochastic and dynamic, many unknown factors may play an important role, making accurate estimation of the vaccine effectiveness a great challenge in both randomized clinical trials and population-based observational studies. To stimulate discussions, we illustrate (incompletely) some challenges and recently developed statistical and machine learning methods for treatment effect estimation as follows.

Methods for clinical trials:

1. Recently, some investigators challenged the view that randomization implies unconfoundedness and claims that randomization and unconfoundedness are two separate concepts. We need further validations on both theoretical and numerical analyses about the relationships between randomization and unconfoundedness.
2. The Cox proportional-hazards model is widely viewed as a causal model and used for treatment effect estimation in both randomized clinical trials and observational studies. However, some researchers concluded that the Cox hazard ratio is not causally interpretable and the Cox model still may suffer selection bias.

Methods for observation data:

1. Estimating the vaccine effectiveness in observational data raises two principal challenges. First, the treatment assignment mechanism is not known a priori. Therefore, there might be confounders, affecting both the vaccine and protection outcomes, which lead to selection bias. The second challenge is censoring. Censoring might be informative, which also may lead to bias.
2. Machine learning (ML) methods for estimating the effects of treatments, including tree-based methods, the nonparametric Random Survival Forest, Bayesian Additive regression trees, Gaussian processes and, in particular, neural networks have grown rapidly. To address the challenges of selection and censoring biases, counterfactual survival analysis has been developed to incorporate censoring and balanced repre-

sentation for individual treatment effect (ITE) prediction. These methods can be applied to COVID-19 vaccine effect analysis and their merits and limitations need to be investigated.

3. To overcome the bias of Cox proportional-hazards model, existing methods for estimating treatment-specific hazard functions are able to address the challenges due to the potential presence of multiple sources of covariate shift: (i) non-randomized treatment assignment (confounding), (ii) informative censoring, and (iii) event-induced covariate shift.
4. The effect of the vaccine may depend on the true patient status. The true patient status, however, is unknown and hidden. In addition, the confounding variables that affect both the potential outcomes and treatments are also hidden. The variational auto-encoder (VAE) that models the hidden variables can be integrated with the Cox model for unbiased treatment effect estimation. VAE-Cox model can be explored for COVID-19 vaccine effect analysis.

We welcome original research articles (methods or computing), case studies, tutorials, reviews, and perspective articles. The formats are flexible. For example, short opinion or commentary articles are welcome regarding misleading discoveries in the literature and potential misuses of statistical or data science methods. Such articles may be as impactful as full-length research articles.

Since 2003, Journal of Data Science has published research works on a wide range of topics that involve understanding and making effective use of field data. The journal has been reformed since July 2020 to better serve the data science community in the era of data science. Attractive features of the journal are completely free access, fast review, and reproducible data science. We look forward to your submissions.

Guest Editors:

- Wenjiang Fu, Department of Mathematics, University of Houston
- Annie Qu, Department of Statistics, University of California, Irvine
- Momiao Xiong, Department of Biostatistics and Data Science, University of Texas Health Science Center at Houston