Editorial: Data Science in Action in Response to the Outbreak of COVID-19 in China

Dean Follmann[1], Peter X. K. Song[2], Hansheng Wang[3], and Jun Yan[4]

[1]*Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA*
[2]*Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA*
[3]*Guanghua School of Management, Peking University, Beijing, China*
[4]*Department of Statistics, University of Connecticut, Storrs, Connecticut, USA*

This special issue contains five articles in Chinese with English abstracts on "Data Science in Action in Response to the Outbreak of COVID-19." It is a sister issue to the special issue of July 2020 on the same theme (Follmann et al., 2020).

As a journal in English, why is *Journal of Data Science* publishing articles in Chinese? When we initiated the call for contributions to the special issue in late February, 2020, China had just experienced the darkest time. We had no idea that the pandemic would become what it became later worldwide. Many statisticians and data scientists in China actively participated in the nationwide fight against the pandemic with their expertises. Some of their works made great impact by providing bases for governmental policies. On March 15, a virtual conference on "Statisticians' Responsibility in Fighting COVID-19 with Science" was hosted by the Capital University of Economics and Business. The one-day conference was well-attended by thousands. It featured 21 presentations, covering a wide range of topics on statistical contributions in response to the pandemic and its aftermaths. We advertised the special issue at the conference. One feedback suggested that we publish articles written in Chinese to rapidly communicate the most recent data science contributions for a timely impact on the practice in China. The suggestion was supported by the Management Board of the journal as a one-time experiment under the circumstances. Therefore, we developed a call for contributions for articles in Chinese, which eventually lead to this special issue.

All articles in this special issue were peer-reviewed with the same criteria as those in the 2020 July issue. Huang et al. (2020b) used masked mobil device data and publicly released epidemic data to identify high-risk carriers of COVID-19 among people who traveled from Wuhan. Visit to any of the designated hospitals in Wuhan and travelers' duration of stay in Wuhan were found to be the most important risk factors. They also estimated number of high-risk carriers that needed to be closely monitored by disease control authorities. This works was presented at the virtual conference in March. Du et al. (2020) studied public sentiment classification using comments on Sina Weibo, the counterpart of Twitter in China, with naive Bayesian models. Their results showed effective reflection of the evolvement of the public sentiment during the pandemic. Ding et al. (2020) reviewed the data release mechanism of official agents and proposed five quality dimensions for releasing COVID-19 data: accuracy, timeliness, systematicness, user-friendliness, and security. They reported quality problems on the official websites of provincial health commissions and suggested improvements. Huang et al. (2020a) studied the impacts of several meteorological factors and air quality indicators on the spread of COVID-19 in two cities, Shenzhen and Wenzhou, in China using distributed lag nonlinear models. Temperature was found to be associated with the number of non-imported confirmed cases in both cities. The nonlinear relationship between the predictors and the transmission of the pandemic had different pattern in the two cities. Liu et al. (2020) assessed the impacts of the public health interventions on the pandemic in the Hubei Province with generalized semi-varying coefficient models for the

number of daily new diagnosed cases. Their results demonstrate that the changes in the real-time transmission rate in Wuhan and other cities in Hubei were almost simultaneous and that the public health interventions had positive effects on reducing the transmission of COVID-19.

We are grateful to the authors for their timely contributions and to the reviewers for their quick actions in reviewing the manuscripts. Mr. Zehan Yang, a gradaute student of the University of Connecticut, did an excellent job typesetting all the articles. His effort is deeply appreciated. Drs. Hansheng Wang and Jun Yan led the editorial task of this issue in Chinese as a subtask of our task force of "Data Science in Action in Response to COVID-19" set out in early Spring 2020 (Follmann et al., 2020). It is our hope that the articles in Chinese has made valuable contributions to the data science actions on fighting COVID-19 in China and worldwide.

References

Ding Y, Zeng Z, Zhang X, Chen A (2020). Quality requirements for the release of COVID-19 data and further regulatory suggestions. *Journal of Data Science*, 18(5): 875–888.

Du X, Huang W, Liu Y, Su H (2020). On public sentiment and topic mining during the COVID-19 pandemic based on sina weibo comment data. *Journal of Data Science*, 18(5): 860–874.

Follmann D, Song PXK, Wang H, Yan J (2020). Editorial: Data science in action in response to the outbreak of COVID-19 in China. *Journal of Data Science*, 18(3): 407–408.

Huang D, Zhao J, Jiang Y, Tian T, Li Y, Wang X (2020a). The impact of meteorological factors and air pollution on the spread of COVID-19. *Journal of Data Science*, 18(5): 889–906.

Huang D, Zhu X, Luo W, Yin H, Hong J, Chen Y, et al. (2020b). On identification of high risk carriers of COVID-19 using masked mobile device data. *Journal of Data Science*, 18: 849–859.

Liu J, Chen Z, Zhang J, Ouyang Y, Guo X, Xu W (2020). The real-time effect of public health interventions on the COVID-19 epidemic in hubei province. *Journal of Data Science*, 18(5): 907–921.