Journal of Data Science 19 (2), 210–214 DOI: 10.6339/21-JDS994E
April 2021

Discussion of "Evaluate the Risk of Resumption of Business for the States of New York, New Jersey and Connecticut via a Pre-Symptomatic and Asymptomatic Transmission Model of COVID-19"

Chuanrong Zhang[1,*] and Xinba Li[2]

[1]Department of Geography & Center for Environmental Sciences and Engineering, University of Connecticut, Storrs, CT, USA
[2]Department of Mathematics & Department of Economics, University of Connecticut, Storrs, CT, USA

1 Introduction

The number of coronavirus cases in the United States has surpassed 8 million by Friday Oct 16, 2020. Based on data from Johns Hopkins University, the case numbers are still steadily increasing daily. Although the Northeast region including New York, New Jersey, and Connecticut has remained relatively stable since the late spring, it is seeing a rise in cases recently. While governors of New York, New Jersey, and Connecticut are working hard to constantly evolve plans to reopen business and schools and relax social distancing restrictions, it may be necessary to slow down the reopening plans as the number of confirmed new cases is increasing with the flu season. It becomes important to evaluate the risk of reopening business for the states of New York, New Jersey, and Connecticut. Tian et al. (2021) makes a timely contribution to provide predicted COVID-19 risk information for the policy-makers to evaluate their reopening plans.

It is known that pre-symptomatic and asymptomatic COVID-19 transmission may make major contributions to the COVID-19 pandemic. Even if all symptomatic cases are isolated, an outbreak may be possible via pre-symptomatic and asymptomatic transmission. However, currently there are few studies to evaluate the risk of pre-symptomatic and asymptomatic transmission of COVID-19 (Slifka and Gao, 2020; He et al., 2020; Gatto et al., 2020), not to say to evaluate the risk of reopening business via a pre-symptomatic and asymptomatic transmission model of COVID-19. Because a large proportion of COVID-19 cases probably come from pre-symptomatic and asymptomatic person transmission, simulated COVID-19 information from a pre-symptomatic and asymptomatic COVID-19 transmission model is needed for policy-makers to make effective public health intervention policies. To safely lift current restrictions and minimize the risk of resurgence and reopening business, symptom based isolation must be complemented by rapidly providing information about the pre-symptomatic and asymptomatic COVID-19 cases. From this aspect, the study of Tian et al. (2021) is also valuable by considering pre-symptomatic and asymptomatic COVID-19 case information which may guide the policy-makers to make better policies for gradual relaxing of community interventions and reopening of business.

2 Challenge Issues

While Tian et al. (2021) provides some valuable contributions to the literature, there are some challenge issues for accurately predicting the COVID-19 situations via the traditional SIR models

*Corresponding author. Email: chuanrong.zhang@uconn.edu.

or the extended SIR models.

The first challenge issue is the data quality issue. It is very difficult, if not impossible, to collect data of infectious individuals with mild or no symptoms who were in-home isolation and recovered. Although some governments made a huge effort to identify suspected cases using different methods such as a door-to-door inspection, it is still very easy to miss some suspected individuals. There is under-reporting of total COVID-19 confirmed cases across the world including New York, New Jersey, and Connecticut. This is because both the clinical cases and the testing capacity are limited. Many individuals with mild symptoms are usually not seeking treatment. The testing capacity is also limited by several factors including resources and training especially during the initial stage of the pandemic. There is a lack of access for testing. Many testing places are only for individuals with obvious symptoms. The limited testing capacity means that we don't know exactly how many people are or have been infected by COVID-19 virus. The public available data including the data reported by governments is certainly missing some COVID-19 cases. The confirmed cases reported by governments probably only represent a partial of the actual infected COVID-19 cases (with or without symptoms). Therefore, the public available COVID-19 data used by the Tian et al. (2021) may contain errors and uncertainty. Data accuracy and incompleteness may cause problems for the accuracy of the model predicted results.

The second issue is the assumption of the extended SIR models. There is no consideration of the spatial heterogeneous issue for many mathematics models of COVID-19 including the extended SIR models implemented by Tian et al. (2021). The study of Tian et al. (2021) has focused on understanding the temporal dynamics of COVID-19 to produce temporal trajectories of the disease under different populations in New York, New Jersey, and Connecticut and didn't assess the spatio-temporal dynamics of the disease. One of the major drawbacks of many SIR class of models, including the proposed models by Tian et al. (2021), is the assumption of well-mixed populations. These traditional or extended SIR models make an assumption that the transmission rate is assumed to be the same for all places within a given state, thus the transmission rate was determined by fitting the state level data to the model. The assumption may not be appropriate given the facts that there is a heterogeneous spatial distribution of the population and the COVID-19 disease across New York, New Jersey, and Connecticut and the mobility of the susceptible individuals and their interaction with infected individuals is also spatially heterogeneous. If the population and infected individuals were spread evenly through the study areas, this assumption would be appropriate and there would be no value in a geographical approach. However, infectious diseases such as COVID-19 have substantial geographic variations in range and intensity of transmission induced by uneven geographic distribution of vulnerable populations and risk factors that facilitate the spatial dispersion of the virus (O'Sullivan et al., 2020; Munshi et al., 2020; Plank et al., 2020; Getz et al., 2019; Brown et al., 2018)

Spatial attributes such as airport and road connectivity were linked to the geographic disparities of COVID-19 related risk of mortality in the U.S (Cuadros et al., 2020). Areas with connectivity enhanced by air transport may have faster spread of COVID-19 infections compared to rural and less connected areas. Understanding the local variation in COVID-19 transmission under heterogeneous geospatial attributes is a crucial step in developing more effective strategies of reopening business. The Geography department at University of Connecticut developed a COVID-19 visualization dashboard for monitoring COVID-19 cases in Connecticut at town level using GIS (Geographic Information System). We have been updating the data daily since March 22, 2020. You can check it out at here: http://COVIDct.org/. Figure 1 shows the spatial distribution of the confirmed COVID-19 cases by Connecticut towns on Oct 16 2020 from the Dashboard. From Figure 1 it can be seen that the distribution of the confirmed COVID-19

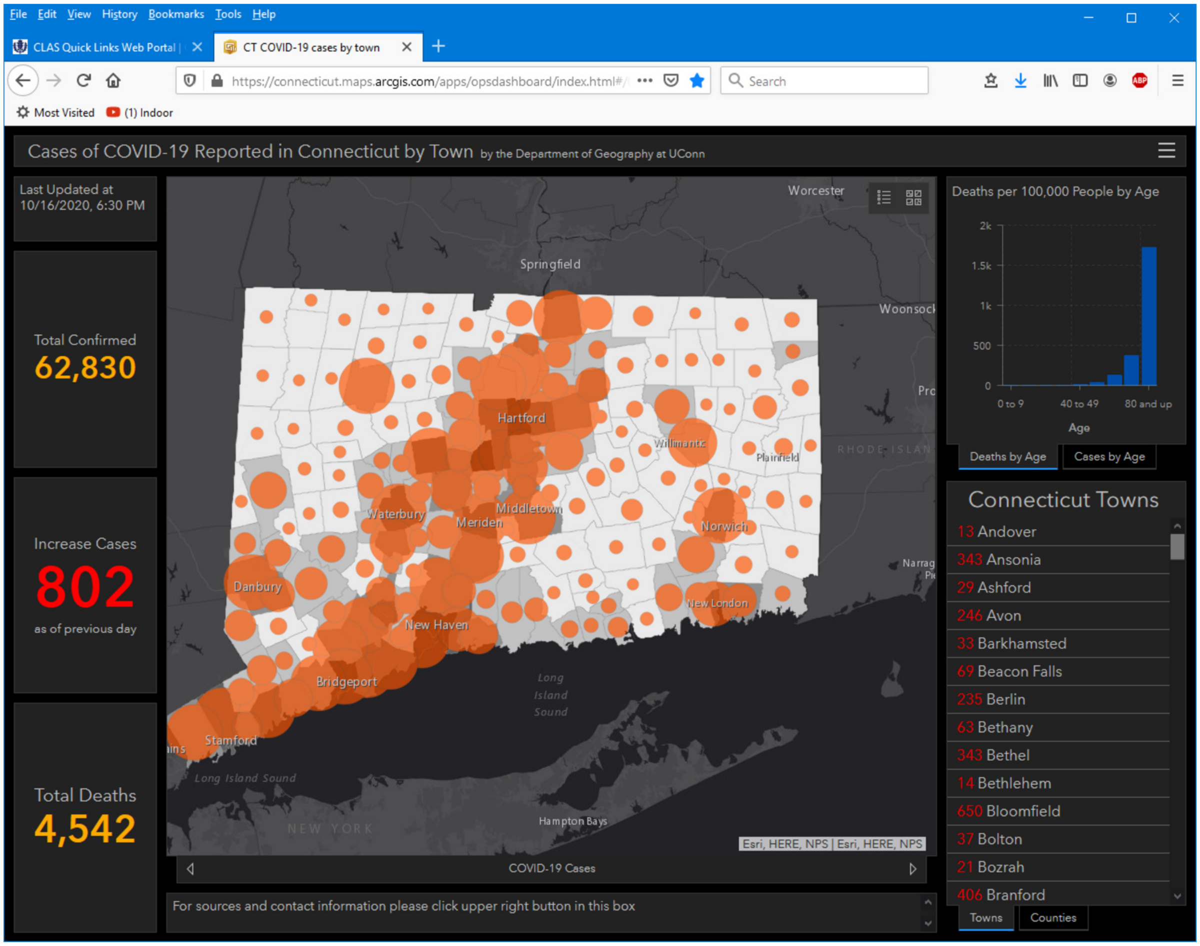

Figure 1: The uneven geographic distribution of the total confirmed cases by Connecticut towns.

cases has an obvious uneven geographic distribution in Connecticut: the urban areas and well connected areas have much more cases than the rural areas and less connected areas. Although there are also data quality issues and probably an unevenly missing cases information in this map as aforementioned (because this map data also come from the publicly available data from multiple sources), this map shows a clear heterogeneity of spatial distribution of COVID-19 cases in Connecticut.

Because the distribution of both population and the COVID-19 disease vary a great deal with place, it is necessary to leverage this when estimating case information for policy-makers to make the reopening business strategy. The spatially-explicit mathematical models are needed to simulate the spatial dynamics of COVID-19 epidemic over discrete time and across discrete space in New York, New Jersey, and Connecticut. The spatially-explicit models may provide useful insights into the epidemiological characteristics of the disease and identification of disease hotspots, which can inform and guide policy-makers for targeted interventions and targeted reopening business in specific locations while limiting state-wide economic disruption across the whole states of New York, New Jersey, and Connecticut. Areas free from infection can be free of restriction as soon as it is deemed safe to do so. The results from spatially-explicit models would make it easier for authorities to locate areas highly affected by COVID-19 and take appropriate intervention actions in that particular areas while open business in other less effected areas across a state.

The third challenge issue is the model uncertainty issue. Because of the data uncertainty issue, the models fitted to the publicly available confirmed COVID-19 data is unlikely to be reliable. The model parameter accuracy is also constrained by our knowledge of the virus. Thus, model parameters associated with COVID-19 transmission also have uncertainty. This parameter uncertainty may propagate through the models, thus unavoidably producing uncertainty in the model results. For example, as aforementioned Tian et al. (2021) assumed that the COVID-19 transmission rates are the same across one whole state (probably for simplification both mathematical analysis and data fitting). In reality, however, the transmission rates may change with different population, socioeconomic status, and outbreak control policies across different places. For some hotspot places, people may also be motivated to take voluntary actions to reduce the contact with the infected individuals and contaminated environment so as to protect themselves and their family members. As a result, the actual transmission rates may decrease even with an ascending outbreak. Of course, we acknowledge that currently there is still a lack of a complete understanding of the factors that affect COVID-19 transmissibility and how COVID-19 transmissibility varies across populations and environment settings especially with the prevalence of pre-symptomatic and asymptomatic infections. However, the independence assumption of the traditional non-spatial SIR models or the extended non-spatial SIR models by Tian et al. (2021) may contain uncertainty due to the existence of spatial autocorrelations in COVID-19 confirmed cases. The independence assumption of these non-spatial SIR models may lead to a biased estimation of standard errors of model parameters, and consequently, mislead significance tests. Instead of showing the spatial variation of the outbreak of COVID-19, these non-spatial SIR models can only provide an "average" of the outbreak in a geographic region such as a whole state.

3 Concluding Remarks

It still remains very difficult to measure and model COVID-19 transmission, especially the pre-symptomatic and asymptomatic transmission under various reopening scenarios. There are many factors related to COVID-19 transmission, ranging from societal and political issues to cultural and ethical standards, which are difficult to be represented in the mathematical models. The models implemented by Tian et al. (2021) may need to incorporate more factors to better simulate the complex reality of COVID-19 situation. However, we must also realize that the increased complexity of a model normally comes with increased difficulty for manipulation, analysis, computation, and implementation. The simpler model issue should not discount the value of the implemented models by Tian et al. (2021).

In general, we acknowledge that it is impossible for any single model to predict the course and situation of the COVID-19 epidemic very accurately. Local responses, time varying strategies and interventions, and other spatio-temporal demographic and environmental factors all play an important role in COVID-19 transmission. Although models as these implemented by Tian et al. (2021) can only serve as a guiding tool to policy-makers for evaluation their reopen business policies, they make a timely contribution by providing important COVID-19 prediction information for effective risk evaluation of reopen business for the states of New York, New Jersey, and Connecticut.

References

Brown GD, Porter AT, Oleson JJ, Hinman JA (2018). Approximate Bayesian computation for spatial SEIR (S) epidemic models. *Spatial and Spatio-Temporal Epidemiology*, 24: 27–37.

Cuadros DF, Xiao Y, Mukandavire Z, Correa-Agudelo E, Hernández A, Kim H, et al. (2020). Spatiotemporal transmission dynamics of the COVID-19 pandemic and its impact on critical healthcare capacity. *Health & Place*, 64: 102404.

Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences*, 117(19): 10484–10491.

Getz WM, Salter R, Mgbara W (2019). Adequacy of SEIR models when epidemics have spatial structure: Ebola in Sierra Leone. *Philosophical Transactions of the Royal Society B*, 374(1775): 20180282.

He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5): 672–675.

Munshi J, Roy I, Balasubramanian G (2020). Spatiotemporal dynamics in demography-sensitive disease transmission: COVID-19 spread in NY as a case study. ArXiv preprint: https://arxiv.org/abs/2005.01001.

O'Sullivan D, Gahegan M, Exeter D, Adams B (2020). Spatially-explicit models for exploring COVID-19 lockdown strategies. *Transactions in GIS*. Forthcoming, https://doi.org/10.1111/tgis.12660.

Plank MJ, Binny RN, Hendy SC, Lustig A, James A, Steyn N (2020). A stochastic model for COVID-19 spread and the effects of alert level 4 in Aotearoa New Zealand. MedRxiv preprint: https://doi.org/10.1101/2020.04.08.20058743.

Slifka MK, Gao L (2020). Is presymptomatic spread a major contributor to COVID-19 transmission? *Nature Medicine*, 26: 1531–1533.

Tian T, Tan J, Jiang Y, Wang X, Zhang H (2021). Evaluate the timing of resumption of business for the states of New York, New Jersey, and California via a pre-symptomatic and asymptomatic transmission model of COVID-19. *Journal of Data Science*, 19. In this issue.