Journal of Travel Medicine, 2020, 1–3
doi: 10.1093/jtm/taaa084
Research Letter

Research Letter

Detecting early signals of COVID-19 global pandemic from network density

Amanda M.Y. Chu, PhD[1], Agnes Tiwari, PhD[2,3], and Mike K.P. So, PhD[4,*]

[1]Department of Social Sciences, The Education University of Hong Kong, Tai Po, Hong Kong, [2]LKS Faculty of Medicine, The University of Hong Kong, Pokfulam Road, Hong Kong, [3]School of Nursing, Hong Kong Sanatorium & Hospital, Hong Kong and [4]Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

*Corresponding author: Prof. Mike K. P. So, Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong; Email: immkpso@ust.hk

Submitted 13 May 2020; Revised 23 May 2020; Editorial Decision 24 May 2020; Accepted 24 May 2020

Key words: Connectedness, coronavirus, disaster preparedness, epidemic risk management, healthcare analytics, population movement, time series

The 2019 novel coronavirus disease (COVID-19) has traveled quickly to almost every corner of the world.[1] One of the important triggers must be the population movement among countries.[2] Correspondingly, a question that may draw researchers' attention is whether the level of connectedness among countries may provide us early signals to detect the global outbreak of the virus. Here we provide some evidence to show that we may have detected the high risk of international spread of COVID-19 as early as late February 2020 by investigating network density of countries,[3] the total number of countries with reported confirmed COVID-19 cases, and the daily COVID-19 confirmed cases worldwide.[4] Our proposed network method using the confirmed cases is different from those methods using human mobility data.[5] Unlike the methods based on the susceptible-infected-recovered dynamics or air-transportation networks,[6,7] the network density in this research assesses the connectedness of different countries' confirmed cases' changes to assess pandemic risk.

Using the data from World Health Organization (WHO)'s situation reports,[4] we examined the connectedness of two countries at a specific time t by calculating their correlation of changes in the number of confirmed COVID-19 cases in the past 14 days (i.e. the selected days include the point of the analysis and its previous 13 days), where we can ensure enough data for the rolling analysis of time series. In this research, the data retrieval started from 21 January to 8 April 2020, and thus 3 February 2020 is the first time point in our analysis. If the correlation of changes in the number of confirmed cases between two countries is greater than 0.5, the two countries are regarded as connected and we refer it as a connection in a pandemic network.[8]

We then calculated the network density, which refers to the ratio of the number of connections with respect to the maximum possible connections among countries. In other words, the higher the network density, the higher the tendency that countries' number of confirmed COVID-19 cases increases together, thus indicating increasing pandemic risk. Below are the summary steps to find the network density in day t.

1. Define X_{it} as the number of confirmed cases of country i in day t. Calculate the daily changes in the square-root of newly confirmed cases, i.e. $Y_{it} = \sqrt{X_{it}} - \sqrt{X_{it-1}}$. Statistically, the 'square-root transformation' is to make the transformed counts more stable.[9]
2. For day t, calculate the correlation of Y_{it} and Y_{jt} for country i and country j using the 14-day data, including the past 13 days and day t.
3. Define a connection if the correlation is > 0.5 and count E_t, the number of connections in the pandemic network in day t.
4. Calculate the network density in day t as $\mathrm{D_t} = \frac{2\mathrm{E_t}}{\mathrm{C_t}(\mathrm{C_t}-1)}$, where $\mathrm{C_t}$ is the number of countries to construct the pandemic network in day t.

We then draw a time series plot of D_t, the network density in day t, and compare its trend with the total number of countries reporting confirmed cases and daily confirmed cases on COVID-19 across the globe within the same time period to identify any patterns and early signals of the outbreak of the virus. Details of the research findings are illustrated in Figure 1.

We clearly see two sharp spikes in the time series plot of network density constructed from worldwide countries' data (the red line) in Figure 1. The first spike of the plot is found in

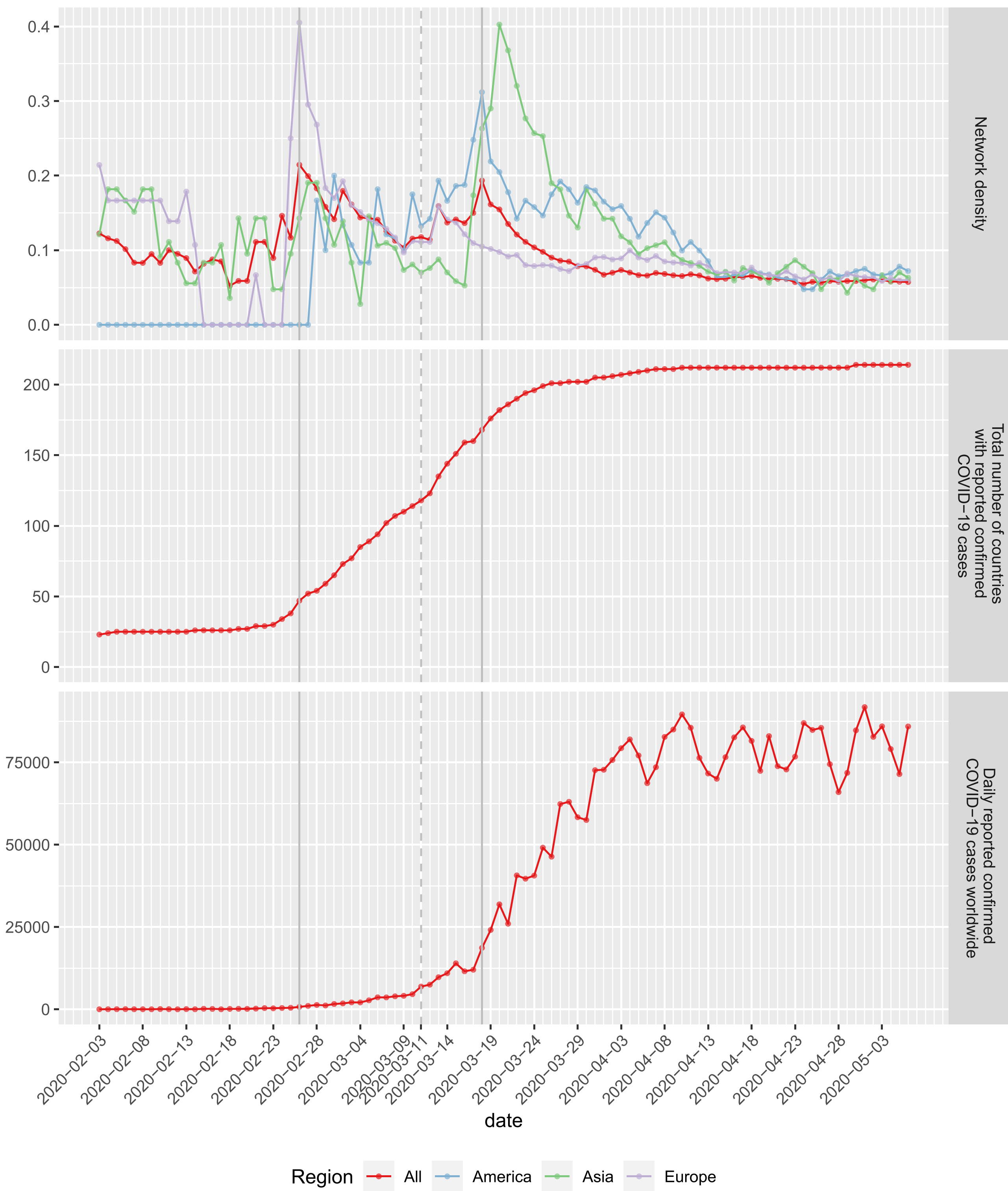

Figure 1. Time series plots of the network density, total number of countries with reported confirmed COVID-19 cases, and daily reported confirmed COVID-19 cases worldwide.

late February (around 26 February), two weeks before WHO declaring COVID-19 a global pandemic on 11 March 2020. Right after reaching the first spike of network density, although the daily reported confirmed cases were at a relatively low level, there was the beginning of a sharp increase in the number of countries with confirmed COVID-19 cases. This could serve as a warning signal of increasing pandemic risk and for countries to consider travel-related issues more seriously and proactively.[10]

The second spike of the plot is found in mid March (around 18 March), a week after WHO announcing the widespread of the virus. After reaching this second spike, not only had the virus affected almost everywhere of the globe, there was also a sudden exponential increase in daily new cases, thus supporting the WHO declaration of the pandemic outbreak. However, the bad news is that by then the spread of COVID-19 was out of control. In future, more attention could be given to time series plots of network density when working on preparedness for pandemic outbreak.

To perform sensitivity analysis, we also calculated the network density based on the data from America (the blue line), Asia (consisting of countries from South-East Asia and Western Pacific defined by WHO, the green line) and Europe (the purple line) in Figure 1. The first peaks of Asia and Europe occurred in late February whereas the first peak of American occurred a few days later. The second peaks of all three network densities fall in mid March. Overall, the patterns of the network density of America, Asia, and Europe are similar to that of the worldwide network density.

We have shown that network density may predict pandemic risk as reflected by the rapid increase in the total number of countries involved in the COVID-19 outbreak and the daily confirmed cases across the globe. In this research, the analysis has been shown to provide the much-needed evidence to signal a pandemic outbreak of COVID-19 as early as late February 2020 and inform timely preparedness for global actions. The network density idea can also be applicable to detect early warning signals of other global pandemics.

Author contributions

AMYC and MKPS conceptualized the study. MKPS collected and analyzed the data. AMYC, AT, and MKPS interpreted the results. AMYC drafted the manuscript. MKPS and AT finalized the manuscript. All authors read and approved the final version of the manuscript.

Source of funding

This work was partially supported by the Hong Kong RGC General Research Fund (grant number 16307217). The funding body had no role in study design, data collection and analysis, preparation of the manuscript, or the decision to publish.

Conflict of interest

The authors have declared no conflicts of interest.

References

1. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med* 2020; **27**(2), taaa021.
2. Watkins J. Preventing a covid-19 pandemic. *BMJ* 2020; **368**:m810
3. Newman M. *Networks*. Oxford: Oxford University Press, 2018.
4. WHO. Coronavirus disease (COVID-2019) situation reports. 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed May 10, 2020).
5. Lai S, Farnham A, Ruktanonchai NW *et al.* Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mHealth for travel medicine. *Journal of Travel Medicine*, 2019; 26(3), taz019.
6. Brockmann D, Helbing D. The hidden geometry of complex network-driven contagion phenomena. *Science* 2013; **342**:1337–42.
7. Wang L, Wu JT. Characterizing the dynamics underlying global spread of epidemics. *Nat Commun* 2018; **9**:218.
8. So MKP, Tiwari A, Chu AMY *et al.* Visualizing COVID-19 pandemic risk through network connectedness. *Int J Infect Dis* 2020; **96**: 558–561.
9. Bartlett MS. The square root transformation in analysis of variance. *Supplement to the Journal of the Royal Statistical Society* 1936; **3**(1), 68–78.
10. Kraemer MUG, Yang CH, Gutierrez B *et al.* The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020; **368**(6490):493–7.