Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Public Health 189 (2020) 3–4

Contents lists available at ScienceDirect

Public Health

journal homepage: www.elsevier.com/locate/puhe

Letter to the Editor

Integrating geographic information system technique with Google Trends data to analyse COVID-19 severity and public interest

At the time of writing, the tally of confirmed novel coronavirus (COVID-19) cases worldwide has exceeded 26.6 million.[1] The United States has become the global epicentre since April 2020, and now it is accounted for nearly one-quarter of the world's total cases. Some studies suggest that health-related issues can cause anxiety which may lead to increased public attention, typically manifested by online information search.[2,3] Along the same lines, given the substantial regional disparities of COVID-19 case severities across states in the United States, the relationship between regional case severities and the public interest emerges as an imperative for COVID-19–based public health studies.

To investigate the relationship between the aforementioned two indicators, geographic information system (GIS) techniques can play a crucial role. Adams et al.'s[4] (2020) GIS-based study points out the shortcomings of using unnormalized COVID-19 demographic data in choropleth mapping, and their use of the normalized data (confirmed cases per 100,000 people) presents a more accurate visualisation of pandemic severity. Although I entirely agree with their point of view and methods, I would like to propose an alternative GIS technique which has the potential to facilitate a better understanding of the research, namely, the cartogram technique.[5,6] A cartogram is a map in which the geometry of areas is distorted to convey the value of an alternative thematic mapping variable.[6] Hence, if the normalized COVID-19–related data is used in a cartogram, it can provide some novel perspectives on data interpretation.

To perform the analysis, the data were obtained from two sources. The COVID-19 case data were retrieved from the US health authority (https://cdc.gov/covid-data-tracker). I retrieved the total confirmed cases per 100,000 population by state, and then I divided

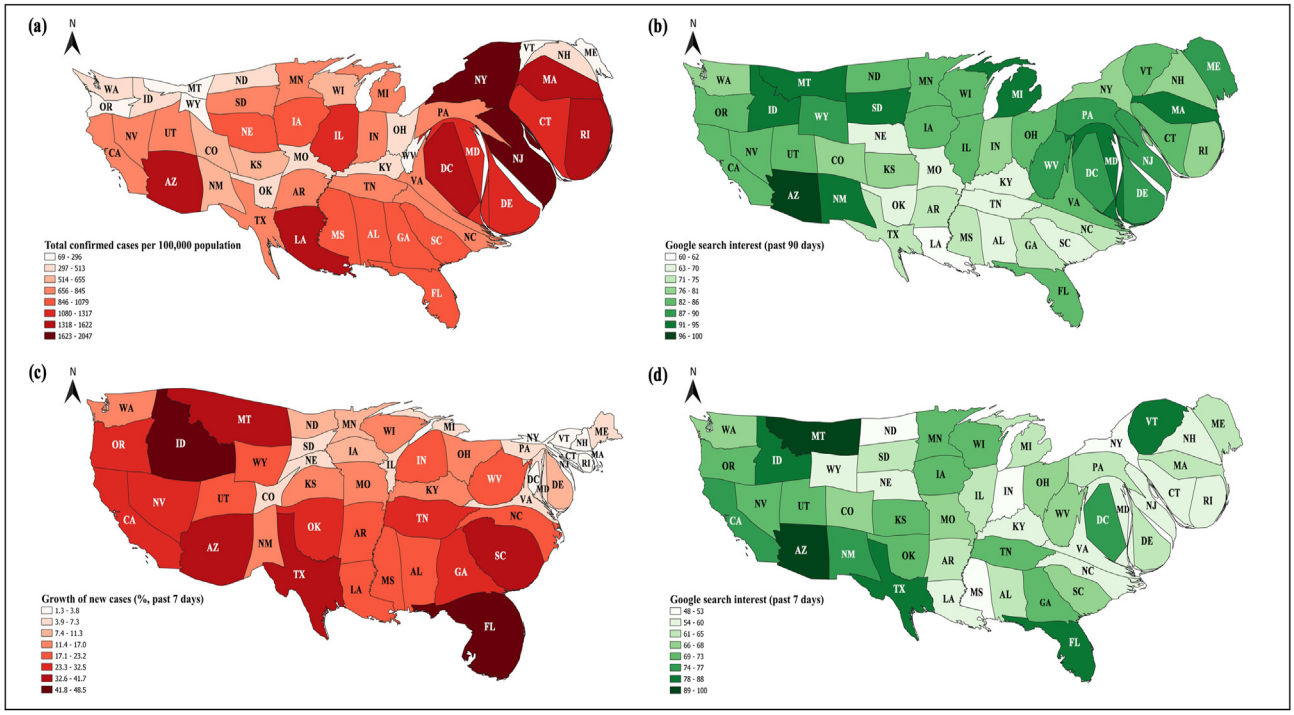

Fig. 1. COVID-19 regional severity and internet search in the United States (as of July 9th, 2020). Data on the cartograms were classified using Jenks Natural Breaks algorithm. The two letters in each polygon stand for the standard ANSI code for the United States state abbreviations. Alaska and Hawaii were included in the statistics but not presented on the cartograms.

https://doi.org/10.1016/j.puhe.2020.09.005

the new confirmed cases (during the past week of data collection) by the total previous cases and obtained a *growth of new cases* indicator. Public interest was captured by people's Google search data in each state.[7] The data were acquired from the *Google Trends* service, which uses a normalized relative search volume comprised between 0 and 100 and thus enables users to compare the search interest of a particular topic across different countries or regions over time. Multiple search terms were tested using the 'related queries' feature. The results revealed that *Coronavirus* was the most searched COVID-19–related term in the United States over the whole period since the outbreak, while *COVID* was the most searched term during the past week of data collection. Therefore, I used the two terms and retrieved each US state's *Google search interest* data (past 90 days and past 7 days, respectively). Using the cartogram technique, I visualised the aforementioned indicators on the US map (Fig. 1).

Fig. 1(a) suggests that the Northeastern region has the highest total case density in the United States. New York and New Jersey are the two worst-hit states, whereas Arizona and Louisiana are the highest outside the Northeast. Comparing it with Fig. 1(b), it shows that the Northeast region states indeed have above-median levels of Google search interest on coronavirus, whereas Arizona takes the national lead. However, the relationship is not significant ($r = 0.07$, $p > 0.05$). For example, Idaho and Montana's case densities are low whereas the search interests are among the highest in the country. Nevertheless, the other two figures can help to untangle the confusion. Fig. 1(c) shows that Idaho and Montana both have increased more than 40% of new cases in the past 7 days and that also lead to spikes in recent internet search. The Northeast region states have stabilised the situation, and most states have less than 5% of case increase; consequently, their search interests drop dramatically. The correlation coefficient between *growth of new cases* and *Google search interest* (past 7 days) is highly significant ($r = 0.65$, $p < 0.001$). The results were in contrast to the findings of Zhang et al.[8] who did not find a significant relationship between internet attention and spread of COVID-19 in China during the early stages. The results also highlight the importance of the quality of internet search in coping with the pandemic which should be further strengthened. Hence, health authorities and internet service providers should strive to provide the most up-to-date and reliable COVID-19 information to the public via the internet.

References

1. Worldometer. *Countries where COVID-19 has spread*. 2020. Available from: www.worldometer.info. [Accessed 5 September 2020].
2. Baumgartner S, Hartmann T. The role of health anxiety in online health information search. *Cyberpsychol, Behav Soc Netw* 2011;**14**(10):613–8. https://doi.org/10.1089/cyber.2010.0425.
3. Lagoe C, Atkin D. Health anxiety in the digital age: an exploration of psychological determinants of online health information seeking. *Comput Hum Behav* 2015;**52**:484–91. https://doi.org/10.1016/j.chb.2015.06.003.
4. Adams A, Chen X, Li W, Zhang C. The disguised pandemic: the importance of data normalization in COVID-19 web mapping. *Publ Health* 2020;**183**:36–7. https://doi.org/10.1016/j.puhe.2020.04.034.
5. Gastner M, Newman M. Diffusion-based method for producing density-equalizing maps. *P Natl Acad Sci* 2004;**101**(20):7499–504. https://doi.org/10.1073/pnas.0400280101.
6. Dorling D. *Area cartograms: their use and creation*. Norwich: University of East Anglia; 1996.
7. Gianfredi V, Bragazzi NL, Mahamid M, Bisharat B, Mahroum N, Amital H, Adawi M. Monitoring public interest toward pertussis outbreaks: an extensive Google Trends – based analysis. *Publ Health* 2018;**165**:9–15. https://doi.org/10.1016/j.puhe.2018.09.001.
8. Zhang H, Chen Y, Gao P, Wu Z. Mapping the changing Internet attention to the spread of coronavirus disease 2019 in China. *Environ Plann A* 2020;**52**(4):691–4. https://doi.org/10.1177/0308518X20922238.

T. Zhang
School of Business and Law, Edith Cowan University, Australia
E-mail address: tenghaoz@our.ecu.edu.au.

4 September 2020
Available online 16 September 2020