

Healthcare statistics based on 'big data' may not always be reliable

December 2 2021, by David Bradley



Credit: Unsplash/CC0 Public Domain

Statistics based on so-called "big data" may not always be as reliable as

we might hope, according to a study published in the *International Journal of Healthcare Technology and Management*. The research analyzed a manageable subset of time-stamped dynamic information from the internet pertinent to COVID-19 infections. Study author Kenneth David Strang of W3-Research in Saint Thomas in US Virgin Islands writes that the results were "surprising" and revealed some limitations to conventional statistical techniques. Strang's work suggests that using general analytics tools for healthcare big data may not be reliable.

Strang points out that while the study is pertinent to our understanding and approach to [big data](#) in the ongoing COVID-19 pandemic, it has broader implications for how big data is analyzed using statistical tools and whether there needs to be a [paradigm shift](#) in our approach and seemingly conflicting ideas that big data can be handled just as we do any other [scientific data](#) or whether such scientific evidence warrants a different approach entirely simply by virtue of the scale of that evidence manifest in big data.

"More research will certainly be needed to verify these reliability problems with healthcare big data since only the coronavirus case study was used here," says Strang. He points out that the nature of big data and a researcher's access to such vast repositories and the [processing power](#) needed to analyze them may offer inherent limitations and how much new information and insight can be readily extracted. Moreover, it is difficult to run checks to prove that any such analysis is valid simply because of the scale of the data and those limitations. Strang offers a hypothetical approach that might allow such validation by using a control data set for a given experiment that is not itself "big" data.

It is almost an aside of the study's findings regarding our approach to big data that Strang was able to demonstrate that there were some "fascinating potential relationships between foreign property ownership

in Australia near the two biggest cities, with links to China, and thereby, potential vulnerabilities to future pandemic outbreaks."

More information: Kenneth David Strang, General analytics limitations with coronavirus healthcare big data, *International Journal of Healthcare Technology and Management* (2021). [DOI: 10.1504/IJHTM.2021.119163](https://doi.org/10.1504/IJHTM.2021.119163)

Provided by Inderscience

Citation: Healthcare statistics based on 'big data' may not always be reliable (2021, December 2) retrieved 24 April 2024 from <https://medicalxpress.com/news/2021-12-healthcare-statistics-based-big-reliable.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.