International Journal of Population Data Science

Journal Website: www.ijpds.org

Infrastructural Inclusion: A method for assessing data infrastructure to advance inclusion, diversity, equity, and accessibility

Amy Freier[1], Morgan Stirling[1], Jeffrey Morgan[2], Laura Bowler[3], and Nathan Nickel[1]

[1]University of Manitoba, Winnipeg, Canada; Health Data Research Network Canada, Vancouver, Canada

[2]University of British Columbia, Vancouver, Canada; Health Data Research Network Canada, Vancouver, Canada

[3]Health Data Research Network Canada, Vancouver, Canada; University of Manitoba, Winnipeg, Canada

Health administration data is often heralded for its ability to reduce bias and allow whole-population inclusion. However, the oft unexamined assumption that administrative data is neutral has perpetuated systemic inequities and enforced policy barriers. Responsibility towards change takes many forms, including within data research organizations and data research infrastructure. Three data infrastructure tools used to conduct population data research were the target of our analysis: i. common data models, ii. concept dictionaries, and iii. phenotype libraries. A literature search was conducted with the aim to understand how principles of inclusion, diversity, equity, and accessibility were considered in their development. Recommendations were made towards future creation of data infrastructure, resulting in a method for infrastructural inclusion. We found that that the academic and grey literature about the development of these tools lacked robust information about IDEA vis-à-vis disaggregated data (sex/gender, race/ethnicity, ability, income), equity in development, and accessibility of data. From these results, we developed specific recommendations for embedding IDEA in common data models, concept dictionaries, and phenotype libraries. More broadly, we developed a tool wherein principles of IDEA can be applied to any future data infrastructure project. Examination of three data research tools helped to identify gaps in consideration for IDEA within technical and infrastructural production of data. A tool for "infrastructural inclusion" shifts responsibility for using large population-based health data equitably from a personal responsibility of individual researchers to a system responsibility of organizations and administrators.

https://doi.org/10.23889/ijpds.v11i5.3608