International Journal of Population Data Science

Journal Website: www.ijpds.org

A Privacy-preserving Aggregation of Distributed Summary-Level Data (PARADISE)

Refik Saskin[1], Donna Curtis Maillet[2], Ted McDonald[2], Robert Platt[3], Colin Dormuth[4], Raquel Duchen[1], Jodi Gatley[1], and Luke Mondor[1]

[1]ICES, Toronto, Canada
[2]DataNB, Fredricton, Canada
[3]McGill University, Montreal, Canada
[4]University of British Columbia, Vancouver, Canada

Small-cell counts – i.e., data-table or stratum cells containing very few individuals – present a persistent disclosure risk in human-research and administrative-data settings. In Canada, layered legal, ethical and statistical disclosure-control frameworks require reasonably robust protections for such data, yet multi-regional data integration poses additional challenges as varying suppression thresholds and linkages increase re-identification risk. Here we introduce a novel privacy-preserving method for combining small-cell data in studies involving three or more data centres in Canada, enabling pooled analysis while respecting provincial and federal confidentiality obligations. We present the methodological framework and application to categorical data, continuous data, and Kaplan-Meier survival data, and discuss implications for data governance, research integrity and public trust. Our findings suggest that the proposed approach strikes a practical balance between protecting individual privacy and preserving analytic value, thus offering a pathway for multi-regional collaboration under Canadian law and ethics guidelines and enhanced analyses of distributed data.

https://doi.org/10.23889/ijpds.v11i5.3583