International Journal of Population Data Science

Journal Website: www.ijpds.org

# Distributed Analytics as a Catalyst for Collaborative Research in Canada

Jean-François Ethier[1], Carrie-Anne Whyte[2], Refik Saskin[3], Erind Dvorani[3], Erik Youngson[4], Xueyi Chen[4], Félix Camirand Lemyre[1], Jean-Philippe Morissette[1], and Kim McGrail[5]

[1]Groupe de recherche interdisciplinaire en informatique de la santé (GRIIS), Sherbrooke, Canada

[2]Canadian Institute for Health Information, Ottawa, Canada; HDRN Canada, Ottawa, Canada

[3]ICES, Toronto, Canada

[4]Alberta SPOR Support Unit, Edmonton, Canada

[5]HDRN Canada, Vancouver, Canada

Health Data Research Network (HDRN) Canada is a distributed network of organizations working together to transform multi-regional health data access and analysis. Central to this strategy is the development and implementation of analytical methods that enable researchers to conduct multi-regional studies within the complexities of Canada's federated data environment, where direct sharing of line-level data across jurisdictions is generally not permitted. Distributed analysis allows statistical analyses to be performed across multiple sites without sharing line-level data. Instead, intermediate privacy-protected numerical results are exchanged between local sites through a coordinating node. This approach respects privacy and data-governance requirements and enables the inclusion of data that cannot otherwise be moved or pooled due to legal, ethical, policy, or social-acceptability constraints. To advance distributed analytics in Canada, HDRN Canada piloted three methods (linear, logistic, and Cox regression) for horizontally distributed data across organizations. Despite demonstrating feasibility, hesitation remains among network members, data custodians, and privacy officers regarding data protection and privacy preservation. Continued assessment and demonstration of minimal risk is essential for strengthening confidence and trust. HDRN Canada is co-developing clear governance mechanisms with its members and data custodians to support responsible implementation and to build the foundation for distributed analytics across Canada. As the network matures, HDRN Canada continues to expand participation and the analytical techniques available. Future directions include piloting vertical distributions, building governance structures to support scalable implementation, and cultivating broader awareness of distributed analytics. Through its collaborations, HDRN Canada is fostering innovation in multi-regional health data use across Canada.

https://doi.org/10.23889/ijpds.v11i5.3613