ECE Data Collection: Canadian Case Studies

Irina Kaliazine and Kerry McCuaig

Atkinson Centre for Society and Child Development, OISE, University of Toronto

INTRODUCTION

- Data linking is the process of creating links between records from different sources based on common features present in those sources. [1]
- Data linking allows for longitudinal analysis as well as assessing the impact of programs and interventions on child and family outcomes.
- In the field of early childhood education (ECE), data on child and program outcomes gives unlimited possibilities to evaluate and rethink public initiatives.
- ECE data supports evidence-based decision-making processes, which can lead to a significant improvement in resource allocation, and potentially, to better child outcomes.
- Most ECE data in Canada is collected by individual organizations and reported separately. Much of the data is not available to researchers.

RESEARCH QUESTIONS

1. What data on children is collected in different jurisdictions across Canada, and whether there are mechanisms to link the data?

2. What are the challenges of data linking?

3. What is the process of data storage (in data repositories) and usage in the provinces that systematically collect and store integrated linkable data on children?

METHODS

1. Jurisdictional scan

- A cross-jurisdiction scan of best practices on the collection and linking of data on children's development and programming was conducted.

- The research included online reviews of the government and program administration websites, as well as follow-up phone interviews with several officials from various jurisdictions.

2. Case study analysis

- Four case studies were completed to investigate Canadian examples of data collection systems with the capacity to link data on child and program outcomes.

FINDINGS

1. Although there are a wide variety of programs and initiatives available for families and children, the study was focused on programs where personal data is collected, and has the potential to be linked. The following categories of programs were identified:

Prenatal			Birth			Infancy	Preschool	School Transition	School	
FAS program	Prenatal Benefit Program	Genetic screening programming	Maternal Child health	Home visiting	Newborn Screening	Immunization Monitoring System	Program Attendance Information	Early Development Instrument (EDI)	Student academic records	Adolescent Health

2. Outside of small scale, time limited research projects there are few Canadian examples of data collection systems with the capacity to link all or even some health, socio-economic, educational, program quality assessment and child outcome findings over time. This review does however provide four models to build on. Case studies were developed for four Canadian provinces: **Manitoba**, **British Columbia**, **New Brunswick**, and **Ontario**. Summary of the findings from the case studies is provided in the table below.

Table 1		Institute for Clinical Evaluative Sciences (ICES)	NB Institute for Research, Data and Training (NB–IRDT)
Location, scope	Location	Ontario	New Brunswick
	Official Date of Inception	1992	2015
	Type of data holdings	· Health · Population and Demographic	Health records, social assistance, justice, training, education (In development)
	# of data sets	73	12, more in development
	Fees charged	Yes for coordination, extraction, data analysis and reporting	
Authority to hold data and conduct research	Type of holding	Data linkage & Data Repository	Data linkage & Data Repository
	Designated research entity	Yes	Yes
	Legislation governing data access for research	Personal Health Information Protection (PHIPA) section 45/PHIPA section 44/ Freedom of Information and Protection of Privacy Act FIPPA section 21 (1e)	Personal Health Information Privacy and Access Act/ (PHIPPA) Right To Information and Protection of Privacy Act (RTIPPA)
	Data ownership	ICES but governed by data sharing agreements	Resides with the data provider and housed via data sharing agreements
	Conducts research	Yes	Yes
Data Linking and Security	Are Identifier Data transferred?	Identifier Data are transferred to ICES. Values are coded and used for linking between data sets	Data with temp identifier added and direct identifiers removed
	Where is the Unique ID used for linking created	Using OHIP number and identifying information to create an individual ICES number (IKN)	NB–IRDT using identifying data to create an encrypted Personal Health Information Number (PHIN)
	Where does linking take place?	ICES	NB-IRDT
	When is data linked?	Data is linked on an ad-hoc basis per project using the IKN attached to content data files for each data set	Per project
Researcher Data Access Conditions & Analysis Information	Where can a researcher access data?	ICES Site/Remote access sites	Secure virtual research environment
	Researcher Data Access Requirements	· Privacy Impact Assessment · Meeting the requirement of respective data sharing agreements	Research Ethics Board approval
	Who conducts analysis?	ICES analysts	Researchers/Students
	Non academic access to data	No	No

Table 1. Comparison of the Ontario's [2] and New Brunswick's [3] data repositories.

Table 2		Manitoba Population Health Data Repository	Population Data BC
Location, scope	Location	Manitoba	British Columbia
	Official Date of Inception	1990	2008
	Type of data holdings	· Health · Education · Population and Demographic · Social Services	· Health · Population and Demographic · Occupational · Education
	# of data sets	72	19
	Fees charged	Yes for data extract preparation, analyst services	Yes for training, coordination of data extracts and data extraction
Authority to hold data and conduct research	Type of holding	Data Repository	Trusted Third Party for Linkage & Data Repository
	Designated research entity	Yes	No
	Legislation governing data access for research	· Personal Health Information Act (PHIA) section 24 – MCHP is a research organization · PHIA section 24 governs disclosure to researchers	· Freedom of Information and Protection of Privacy Act (FIPPA) section 33 in accordance with section 35
	Data ownership	Resides with the data provider and housed via data sharing agreements	Resides with the Data Steward as a public body and housed via data sharing and information sharing agreements
	Conducts research	Yes	No
Data Linking and Security	Are Identifier Data transferred?	Identifier Data are transferred to Manitoba Health. It creates an encrypted PHIN (unique ID). This is transferred to MCHP for linking between datasets.	Identifier Data are transferred to Population Data BC for creation of a unique ID for linking between data sets.
	Where is the Unique ID used for linking created	Manitoba Health using identifying data to create an encrypted Personal Health Identification Number (PHIN)	Population Data BC using identifying information to create PopData ID
	Where does linking take place?	MCHP	Population Data BC
	When is data linked?	Data is linked on an ad-hoc basis using the encrypted PHIN attached to the content data file for each data set	Data is linked on an ad-hoc basis using the PopDataID attached to the content data file for each data set
Researcher Data Access Conditions & Analysis Information	Where can a researcher access data?	· MCHP Site · Remote access sites (9)	· Population Data BC site · Remote access via VPN
	Researcher Data Access Requirements	· Signed UM–MCHP researcher agreements · Manitoba Health Information Privacy Committee (HIPC) and UM Health Research Ethics Board (HREB) · Completion of Accreditation Process · Pledge of confidentiality	· Ethics Approval · Scientific merit · Signed Research Agreement with Data Steward(s) · Signed Confidentiality Agreement · Privacy Training
	Who conducts analysis?	· MCHP analysts · Identified external analyst(s)	· Researchers · Named team members
	Non academic access to data	Project may be funded externally however tenured researcher leads the project and all analysis must be conducted by MCHP staff and students	Not available at this time.

Table 2. Comparison of the Manitoba's [4] and British Columbia's [5] data repositories.

DISCUSSION

- In order to obtain more comprehensive information on program efficacy and child outcomes, it is critical to improve systematic data collection and analysis.

- Unlinked data considerably limits possible ways to analyze and use this data.

- Data collection and reporting are essential to evidence-based decision making and democratic accountability.

- **Future Discussion: How can datasets be linked and analyzed while preserving individual privacy?**

REFERENCES

[1] National Statistical Service (2018). *Statistical Data Integration*. Retrieved from http://www.nss.gov.au/nss/home.NSF/pages/Data+Integration+-+Glossary.
[2] Institute for Clinical Evaluative Sciences (2018). *ICES Data*. Retrieved from https://www.ices.on.ca/Data-and-Privacy/ICES-data
[3] University of New Brunswick (2018). *NB Institute for Research, Data and Training*. Retrieved from http://www.unb.ca/fredericton/arts/nbirdt/services-researchers.html
[4] University of Manitoba (2017). *Manitoba Population Research Data Repository – Overview*. Retrieved from http://www.umanitoba.ca/faculties/health_sciences/medicine/units/community_health_sciences/departmental_units/mchp/resources/repository/index.html
[5] Population Data BC (2018). *Population Data BC*. Retrieved from https://www.popdata.bc.ca/