Integrated Data Repositories for Child Development in Atlantic Canada

A Progress Report

August 2020

Dr. David Philpott, Ed.D.

---

www.davidphilpott.ca
david@davidphilpott.ca

with the support of the Margaret & Wallace McCain Family Foundation Inc.

Introduction

Data drives public policy. It informs programs and maps populations. It validates intervention. It monitors need and raises alerts. It makes causal connections. Data is the evidence on which public investments are built.

In the digital era, data is rich, plentiful and current. Cellphones map our movements and computer caches track our interests. Electronic medical records monitor our health while digital pharmacies track our wellness. We live in a world where we are entrants in multiple data sets. The potential for this to inform our collective well-being is endless.

Increasingly, governments are looking to the wealth of data that is currently collected on children to inform public policy. Early identification and intervention is the mantra of child development, and the importance of the early years in optimizing human potential is exceptionally well documented.[i] Today, *earlier* identification and intervention begins in the first five years of life. Using standardized tests to screen for developmental risk is fraught with concerns. While numerous testing instruments are marketed as screens for young children there is near universal concern for the accuracy of such measures[ii]. The uneven nature of 5-year-olds and younger children often result in low test scores that reflect prior experience rather than developmental ability. Too many false positives can cause emotional distress to families and result in significant pressure to initiate expensive interventions that may not be necessary. Such tests tend to focus on deficits at the expense of indicating strengths which will only raise undue concern rather than accurately inform programs. Finally, these instruments are point-in-time measures and run the added risk of not being able to comment on the effectiveness of intervention, necessitating a re-test at some point. While the intention of catching vulnerable children early is commendable, universal access to quality early learning, coupled with efficient documentation and communication, not only monitors individual need and developmental trajectory, it also stands to inform program development. Educators and parents know their children. Capturing that knowledge so as to inform programs and empower families is crucial.

Integrated data repositories hold the potential to realize these goals. Government systems and departments collect a vast amount of data on children beginning with perinatal information on maternal health to APGAR scores, public health records, electronic medical files and doctor billing profiles, pharmacy networks, and early child education programs. This data mushrooms at school start when electronic datasets such as Power School record everything from home address, attendance records and academic, social, and behavioral development. While systems are honing the collection of data, there has been less effort on finding ways to interpret data collectively and provide a comprehensive profile on a child, a family and a community. However, the aggregation of data from multiple sources and perspectives allows for multiple lines of inquiry that can produce textured profiles of neighborhood, communities and regions. Allowing rich data analysis allows researchers to track large or small groups of children, exploring their outcomes along multiple data points and then examine impact of interventions. For example, school attendance patterns could be tracked and analyzed by participation in early years programs, academic performance, and behavioral regulation as indicated on school reports. This could be further broken down by postal code, ethnicity, and prescriptions for ADHD medication, yielding helpful and current information on the developmental needs of children as well as the impact of very specific interventions. At any given point in time repositories have multiple studies ongoing, examining full populations, smaller regional groups by certain demographic markers and highly specific groups by individualized markers. Aggregated data from rich, multiple perspectives stands to not only track human development but to inform public policy and guide the efficient delivery of interventions.

This project sought to explore where the Atlantic provinces are in the development of data repositories to monitor child development. While each province is working toward this goal, they are at various points in the journey and, as such, stand to help inform one another. The project sought to begin a conversation among the provinces to share their valued experiences and knowledge. It was hoped that these conversations would not only speed their progress but also ensure opportunities for collaborative professional learning and research.

Methodology

A series of interviews with key informants in each of the Atlantic provinces was conducted in the spring of 2020. Websites were explored and a brief literature scan was also conducted. It was not meant to be expansive or in-depth but rather a point in time exploration of where each Atlantic province is: what is common, what is unique and what opportunities exist to promote valuable collaboration? A draft report was developed and forwarded to the key informants for triangulate the data before the final report was completed.

The information is presented as emergent themes - “food for thought” for those engaged in the work and interested in the use of data repositories. The project ends with suggested next steps that emerged during the work.

Background

Integrated data systems (IDS) have emerged in today’s electronic era as the result of service providers digitally recording data on the clients they serve and then reviewing that data to ensure effective and efficient public services[iii]. While there is great variety in what is considered an IDS, most have a number of common components. First, an IDS contains data collected during routine practices of operating education, health, justice and social service programs. Among these metrics is demographical information on the client which can include indicators of their well-being. A second common feature is the system’s ability to link data across multiple systems of service and thereby track and monitor their development. To do this, a unique identifier is required to facilitate the matching of individual records. An IDS also has the ability to review the data over time and thereby comment on the impact of earlier experiences and explore multiple lines of inquiry[iv]. A fourth feature is a metadata repository function with information on data quality (e.g., rates of missing values)[v]. Finally, a fifth feature of IDS is a governance system that manages the architecture of the system while ensuring confidentiality of the data and privacy legislation for access. Governance of an IDS assumes responsibility for policies aimed at research and not clinical care, including:

- Data use and access, including data sharing, interpretation and training
- Compliance risk such as monitoring, measuring, and controlling the disclosure of sensitive client information
- Data quality monitoring to ensure accuracy, consistency, and reliability of data
- Technology architecture and system maintenance
- Clarity of the roles and responsibilities of data providers and consumer agencies for each data set represented within the IDS[vi].

In addition to research opportunities, IDS play vital roles in service provision and public policy. Notable among these roles are:

- Policy analysis gleaned from insights into client populations and service patterns
- Identification and stratification of target populations, especially with cross agency access
- Utilization analysis which can help allocate interventions and services more effectively
- Assessment of experiences/development and monitoring of interventions and outcomes
- Service coordination to address gaps and overlaps
- Cross-agency alerts to notify of significant events and emergent needs[vii]

Using IDS to track child development over time and across multiple lines of inquiry is hardly new. Longitudinal studies on early child education have been ongoing for the last 60 years. In the US, the National Governors Association Taskforce on School Readiness (2005) recommended that each state invest sufficient resources to train professionals and establish the mechanisms to share and analyze relevant data beginning with the early years and continuing into school[viii].

In Canada, IDS are becoming increasingly popular. A 2018 environmental scan[ix] noted four provinces with some version of a repository in operation, including British Columbia[x], Ontario[xi], Manitoba[xii] and New Brunswick[xiii], with Manitoba's operating since 1990 and New Brunswick's being relatively new (2015). Manitoba reports that it collects de-identified information that allows it to follow individual interactions without identifying the person. The repository is open to researchers, academics and government staff through a rigorous application and screening process. Housed at the University of Manitoba, it collects data from a number of government departments including:

- Health care - visits to hospitals, physicians, emergency departments and specialists, homecare, and personal care homes, and pharmaceutical prescriptions
- Education - school programs, grades, graduation rates, programs to support children, and adult training programs to support employment
- Social/Families – services provided to support families, children in care, employment/income assistance and social housing
- Justice – prosecutions, court appearance, and charges/outcomes at disposition

- Registries - the Manitoba Health Insurance Registry and Vital Statistics Mortality Registry[xiv]

Atlantic Canada Progress - Emergent Themes

Conversations with key stakeholders, with a perspective of child development, involved in establishing repositories in Atlantic Canada yielded the following themes.

1. ***Growing interest.*** People were eager to engage in conversations on the progress they are making in establishing a repository. They spoke with excitement about the potential for research and were well aware of other models such as MB's and NB's, speaking with envy about the benefit that it could bring to their province. Most acknowledged that it has been a difficult sell to convince government of the importance of this initiative but outlined that people are increasingly supportive of the idea and see the critical role it will play in monitoring need and informing interventions and policy. Participants see the importance of monitoring child development from an early age so as to support their development more effectively.
2. ***Support from the premier.*** NB and NL credited their rapid progress as stemming from a directive from the premier to mandate all government departments to collaborate in establishing this. With that mandate departments have worked collegially and rather quickly to move the process along. PEI and NS do not have a directive from the premier and collaborating with government departments is much more cumbersome. NS does have a government initiative to establish "one patient, one record" which will streamline case planning and reduce duplication. PEI has an initiative to look at the "first 1000 days of a child's life" to explore data collection and service use. Both these initiatives are helping create interdepartmental sharing and planning that could expedite the establishment of a repository.
3. ***Cross-departmental collaboration.*** "Departmental silos" surfaced continuously in conversations and was voiced as the single biggest obstacle to progress. Apprehension on information sharing, confusion of roles, duplications of services, budget limitations and "turf concern" were all identified as reasons why a repository was needed so as to ensure effective and efficient data collection and interpretation.
4. ***External data management group.*** Each province envisions the repository as being held in a structure external to government. NB's is housed at the University of New Brunswick (NB Institute for Research Data and Technology). NL's will be housed at the NL Center for Health Information[xv]. PEI are in early conversations with two possible hosts while NS is also in conversations with Maritime SPOR Support Unit[xvi]. Participants outlined the need for the repository to be well regulated by government but managed independently.
5. **Information sharing and privacy concerns.** The need to ensure confidentiality was paramount and participants were well aware of the progress that NB has made in passing legislation to modify the Personal Health Information Act to allow the sharing of information and new legislation to guide the establishment and operation of the

repository. NL, PEI and NS all voiced a desire to learn from NB's experience and staff at NB-IRDT voiced a willingness to provide this professional learning.

6. ***Ethical review and stringent research requirements.*** The need to establish application processes, including ethical review clearance, to access the data set was evident. Although a repository has untold potential for research, the confidential nature of working with human subjects was evident and requires careful safeguarding.
7. ***Reliance on medical records as common data platform.*** Each province reported that data collection begins with medical information and grows from there. Health data is a lifelong, consistent data source that can be complimented by other data. There was interest in collecting maternal health data and then tracking forward with APGAR scores, healthy baby checks, vaccinations, public health records, pre-school checks, medical records, referrals, prescriptions, etc. NB and NL are currently using electronic medical records and pharmacy networks while PEI and NS are in the process of working towards this. Those two provinces currently use doctor billing codes. NL cautioned that they are phasing in electronic medical records and not all physicians are there yet.
8. ***Unique birth identifiers.*** While medical records are the starting point of data collection there was universal apprehension of using provincial health numbers to code the data. NB encrypts a special code from the health number to create a unique identifier that allows the data to be pseudonymous. They are working towards one birth identifier for all government departments and services. All participants see a unique birth identifier as being imperative to privacy.
9. ***Clean data.*** Participants reported being engaged in trying to determine what data is currently being collected in the province, as well as how and when it is being collected to ensure that the data is clean and comparable. NL reported a need to clearly define what data is to be collected and how this is to be reported using a provincial approach to data collection. NS reported much less regionalization due to the presence of only two authorities.
10. ***Scant data on early child education (ECE).*** Excepting data on young children with identified exceptionalities, each province reported little data on enrollment, participation rates and progress in early childhood education programs. The privatized nature of ECE programs and the absence of a common birth identifier were seen as obstacles to collecting such data despite a great interest in being able to do so. Participants felt that being able to track participation and enrollment (full or part time), duration and quality of the program would all be extremely helpful data to have. PEI reported that they have an opportunity to collect this data with their Chances and affiliate programs which deal with most of the province's young children. The insights of early child educators on each child's development and progress is widely seen as being extremely valuable and a missed opportunity for the secondary system. Greater transition planning was universally called for.
11. ***School start - the data mushroom.*** The electronic data system used by schools, such as Power School, amasses a wealth of data on child development. Supplementary systems, such as Review 360, hold the ability to collect even more, especially on children with additional learning needs. When matched with health data and demographics such as postal code, the research potential also mushrooms.

12. ***Linking provincial and national data.*** Postal codes, income support programs, employment status of parents, immigration landing data, census data, etc. afford the opportunity for population health and community profiling. As children age and move into post-secondary schools and eventually the workforce, longitudinal research opportunities arise. Long-term impacts of early experiences (positive and negative) and causal factors can easily be measured and reported on.
13. ***Population health profiles at 5 years of age.*** The continued use of the Early Development Instrument (EDI) was discussed as providing a population level profile of the developmental needs of young children. Efforts have been made to coordinate administration across the provinces to allow a regional analysis. PEI and NL completed the EDI in 2019 and NS completed it in 2020 (NS has EDI data for 2013, 2015 & 2018). NB has cancelled administration of EDI but they, along with PEI, use the Early Years Evaluation – Teacher Assessment (EYE). Participants referenced including the EDI in the data repositories to provide an overview of all children and developmental trajectories when compared to previous administrations. Education Ministries have the capability of matching EDI data with other academic outcomes (standardized testing, report cards, individual supports, attendance, etc). NS uses the EDI as a monitoring tool as well as to inform professional development, resource allocation and program planning in public health, early intervention, and at the school and district level. Neither NB nor PEI could outline how the results of the EYE are being used in their provinces, other than grade composition. The EYE is added to NB's repository but it has not been analyzed. There was some discussion that with a rich and diverse repository, with the potential to track children by postal code, instruments such as the EDI and the EYE could be less helpful in the future. At the same time, concern was also expressed that without the EDI data as a universal touch point, developmental trajectories would have a significant gap between early health indicators to test scores in elementary school.
14. ***Dwindling interest in testing toddlers.*** Each province referenced some interest in screening toddlers for developmental risk. PEI discussed using Ages and Stages but have held back at 18-month immunization conducted by public health. NL has considered the Toddler Development Inventory and NS has considered the Early Years Check-in. Interest in these measures was more casual with apprehension persisting of cost-benefit analysis. Most felt that capturing the perspectives of early child educators and parents would be a more accurate and helpful method of screening young children rather than obtaining test scores, whatever the instrument. NS is working towards a common "report card" for early years programs as a step toward authentic assessment of the curriculum framework, including observational guides for younger children. The intention is to capture the valuable perspectives of early educators so as to facilitate early intervention. Central to this discussion was the need for better transitioning of this information through to school start.
15. ***Identifying key markers.*** Participants were optimistic about the potential of repositories to provide powerful, current data on the exact needs of the populations of young children they are serving. High needs populations, such as children with exceptionalities and children in care, are more easily identifiable and (trackable). However, the prospect of having detailed information on the needs and evolving developmental trajectories of

all children is more complicated. They voiced possibilities of what data could be collected and how to start this, given the current dearth of data on early years participation. Markers such as enrollments/attendance, duration and quality of the program, and training of staff were all seen as information that would be helpful. There was also discussion about how to capture the observations of educators as well as children's progress with meeting curriculum goals – such as NS's efforts for an early years progress report. As one participant stated, "we need a way to quantify the invaluable qualitative observations that educators gather on these children's development". Of particular interest was children's language skills, social interactions and behavioral regulation. Again, transition planning was central to this discussion.

16. ***Opportunity for an Atlantic Canada repository.*** Finally, there was interest in the opportunity for data sharing and creating the opportunity for broader analysis on all children in Atlantic Canada if the repositories are structured in a way to code similar data points. Each province faces similar population challenges – dwindling communities, urban/rural divide driving supply/demand, limited resources, growing international population, challenges with attracting trained early child educators, and struggles to link with the secondary system. Research findings that emerge from a larger study offers greater impact while also informing local need. There was recognition that in order to accomplish this goal greater collaboration would be required during the developmental phases of building these repositories.

Next Steps

The final theme outlined above validates the need for collaboration and information sharing as the four Atlantic provinces continue to move toward the establishment of integrated data systems to track child development and inform public programs. The region has much in common and the opportunities for mutual support are rich. Collaboration in design, collaboration in professional learning and collaboration in research will lead to textured research that will inform public policy and optimize human development.

NB has much to offer the other three provinces with leadership in realizing their goals. A first step is sharing the path towards legislative change that allows information sharing while ensuring personal privacy protection and the effective and efficient operation of a repository. The staff at NB-IRDT have already committed to this and the interest among the provinces to participate is evident.

Ensuring clean data is not as simple as it might appear. Significant professional learning, across multiple government departments, will be required to allow data to be comparable, both within the province and across the region. Central to this will be questions surrounding what data is currently collected, how/when is it collected, and how helpful it is to actually collect the data? A second set of questions will entail what data can be added to this collection? While those questions might be simpler to answer in health, social services and education sectors, significant conversations and planning will be required to answer for the early years sector, currently devoid of data. However, they are critical questions to explore, given the demand for space and the international research on the benefits of quality ECE on human development. Capturing that information will not only allow for in-depth longitudinal research but it will also inform public policy and resource allocations.

Finally, data repositories require the full support of government, emanating from the Office of the premier. A culture of siloes dominate government bureaucracies, which is a significant obstacle to both the establishment and efficient operation of an integrated data system. Leadership from the Premier mandates' collaboration. Collaboration creates efficiency.

References

Alexander, C., Beckman, K., Macdonald, A., Renner, C. & Stewart, M. (2017). *Ready for Life: A Socio-Economic Analysis of Early Childhood Education and Care.* Ottawa: The Conference Board of Canada.

Bornfreund, L., & Severns, M. (2010). Many missing pieces: The difficult task of linking early childhood data and school-based data systems. New America Foundation. Retrieved from https:// static.newamerica.org/attachments/2357-many-missing-pieces/ NAF_ManyMissingPieces.311702b4e89b4b09ae818c049a8797c7. pdf

Burwell, S. (2014). Guidance for providing and using administrative data for statistical purposes (M-14-06). Retrieved from http://www.whitehouse.gov/sites/default/files/omb/ memoranda/2014/m-14-06.pdf

Canadian Taskforce on Preventive Health Care (2016). Recommendations on screening for developmental delay. Canadian Medical Association Journal. 188(8). Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4868607/

Kaliazine, I. & McQuaig, K. (2018). ECE Data Collection: Canadian Case Studies. Retireved from: https://www.oise.utoronto.ca/atkinson/UserFiles/File/Events/20180601_Summer_Institute_2018/Presentations/Summer_Institute_-_Poster_Presentation_-_Kaliazine.pdf

Kumar, P. (2011). An Overview of Architectures and Techniques for Integrated Data Systems Implementation. Actionable Intelligence for Social Policy (AISP), University of Pennsylvania. Retrieved from: https://1slo241vnt3j2dn45s1y90db-wpengine.netdna-ssl.com/wp-content/uploads/2015/09/0033_12_SP2_Architectures_Techniques_Data_Systems_000.pdf

Limlingan, M.C., Grindal, T., Lopez, M., Blocklin, M. & Bymgarner, E. (2015). Integrated Data Systems: An Emerging Tool to Support Services for Low-Income Hispanic Families with Young Children. National Research center on Hispanic Children and families. Retrieved: https://www.childtrends.org/wp-content/uploads/2015/08/IDS-Brief.pdf

MacKenzie, S.L., Wyatt, M.C., Schuff, R., Tenenbaum, J.D. & Anderson, N. (2012). Practices and perspectives on building integrated data repositories: results from a 2010 CTSA survey. Journal of the American Medical Information Association. doi:10.1136/amiajnl-2011-000508

McCain, Honourable Margaret Norrie (2020). Early Years Study 4: Thriving Kids, Thriving Society. Toronto: Margaret and Wallace McCain Family Foundation Inc. Retrieved: https://earlyyearsstudy.ca/

National Governors Association (2005). Building the Foundations for Bright Futures. Final Report of the NGA Taskforce on School Readiness. Retrieved from: https://www.researchconnections.org/childcare/resources/5130

Key Informants

Baker-Worthman, Lisa
Barnes, Eldred
Beer, Janet
Frizzell, Katelyn
Gaskin, Ashley
Gervais, Nicole
Goss-Prowse, Mary
Janus, Magdalena
Jeffers, Elizabeth
Kuhle, Stefan
Leonard, Philip
McDonald, Ted
Montelpare, Bill
Simpson, Carolyn
Smith, Anne-Marie

Endnotes

---

[i] Alexander et al. 2017; McCain, 2020.
[ii] Canadian Taskforce on Preventive Health Care, 2016
[iii] Limlingan, 2015
[iv] Bronfreund, 2010
[v] McKenzie, 2012
[vi] Kumar, 2011
[vii] IBID
[viii] National Governors Association, 2005
[ix] Kaliaziner & McQuaig, 2018
[x] https://www.popdata.bc.ca/
[xi] https://www.ices.on.ca/Data-and-Privacy/ICES-data
[xii] http://www.umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/resources/repository/
[xiii] https://www.unb.ca/nbirdt/
[xiv] http://www.umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/resources/repository/
[xv] https://www.nlchi.nl.ca/
[xvi] http://www.spor-maritime-srap.ca/