



---

## Data Resource Profile

# Data Resource Profile: Pathways to Health and Social Equity for Children (PATHS Equity for Children)

**Nathan C Nickel,<sup>1,2\*</sup> Dan G Chateau,<sup>1,2</sup> Patricia J Martens,<sup>1,2</sup> Marni D Brownell,<sup>1,2</sup> Alan Katz,<sup>1,2,3</sup> Elaine MJ Burland,<sup>1</sup> Randy Walld,<sup>1</sup> Mingming Hu,<sup>1</sup> Carole R Taylor,<sup>1</sup> Joykrishna Sarkar<sup>1</sup> and Chun Yan Goh;<sup>1</sup> the PATHS Equity Team<sup>1</sup>**

<sup>1</sup>Manitoba Centre for Health Policy, <sup>2</sup>Department of Community Health Sciences and <sup>3</sup>Department of Family Medicine, Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

\*Corresponding author. 408-727 McDermot Ave, Winnipeg, MB, Canada R3E 3P5.

E-mail: Nathan\_Nickel@cpe.umanitoba.ca

Accepted 20 August 2014

## Abstract

The PATHS Data Resource is a unique database comprising data that follow individuals from the prenatal period to adulthood. The PATHS Resource was developed for conducting longitudinal epidemiological research into child health and health equity. It contains individual-level data on health, socioeconomic status, social services and education. Individuals' data are linkable across these domains, allowing researchers to follow children through childhood and across a variety of sectors. PATHS includes nearly all individuals that were born between 1984 and 2012 and registered with Manitoba's universal health insurance programme at some point during childhood. All PATHS data are anonymized. Key concepts, definitions and algorithms necessary to work with the PATHS Resource are freely accessible online and an interactive forum is available to new researchers working with these data. The PATHS Resource is one of the richest and most complete databases assembled for conducting longitudinal epidemiological research, incorporating many variables that address the social determinants of health and health equity. Interested researchers are encouraged to contact [mchp\_access@cpe.umanitoba.ca] to obtain access to PATHS to use in their own programmes of research.

### Key Messages

- The PATHS Resource is a population-based suite of numerous data files comprising individuals aged 0–18 years in Manitoba and born 1984–2012 (n = 584 255), capturing information on marginalized populations often lost to follow-up in survey data.
- Individuals' data are linkable across several domains including health, education and social services, allowing researchers to (i) develop a holistic picture of the determinants of health and (ii) model child health/development trajectories.
- The PATHS Resource allows researchers to link data within families, allowing for sibling fixed-effects analyses and studies into intergenerational health disparities.
- More information about the PATHS Data Resource, including how external investigators can apply for access, can be found at [[http://www.umanitoba.ca/faculties/medicine/units/community\\_health\\_sciences/departmental\\_units/mchp/protocol/media/PATHS\\_Resource\\_Website.pdf](http://www.umanitoba.ca/faculties/medicine/units/community_health_sciences/departmental_units/mchp/protocol/media/PATHS_Resource_Website.pdf)].

## Data Resource Basics

### Background of the PATHS Data Resource

Many health outcomes follow a social gradient where lower socioeconomic status is associated with worse health.<sup>1–7</sup> These health inequities have negative societal impacts including increased healthcare expenditures and lost productivity.<sup>8–11</sup> Child health inequities, specifically, can have costs.<sup>12,13</sup>

One can envision a pathway through childhood where opportunities exist for interventions that (i) improve health and developmental outcomes and (ii) reduce health inequities.<sup>14,15</sup> Such interventions may prevent the accumulation of suboptimal outcomes, reduce health/developmental inequities in children and place children on the pathway towards adult health and well-being.<sup>16</sup> Research is required to test the impact interventions have on child health and development and on health inequities. Data are needed that both include children from across the socioeconomic gradient and that follow children through childhood.

### Purpose and scope of the PATHS Data Resource

The goal of the Pathways to Health and Social Equity for Children (PATHS) Data Resource is to provide researchers with comprehensive data to conduct child health and health equity research. The PATHS Resource contains data on children from across the socioeconomic gradient, covering several domains including health, social services and education. It includes clinical, administrative and programmatic data that are linkable at the person level. Data linkage provides researchers with a powerful opportunity to conduct cross-cutting epidemiological research.<sup>10,17–19</sup> Appendix A (available as [Supplementary data](#) at *IJE* online) presents a summary of the databases available in the PATHS Resource.

The PATHS Resource was specifically created to measure the impact that interventions have on child health equity in Manitoba, a province in central Canada with a population of about 1.2 million people. It was generated from the Population Health Research Data Repository at the Manitoba Centre for Health Policy (MCHP). The Repository is a data warehouse where several unlinked, disparate databases reside. Repository data can be used to conduct health research. However, because Repository data are raw and unlinked, one-off research projects require a lengthy and expensive process to both (i) assemble into key indicators and (ii) link at the individual level. Only after completing this extensive process are the disparate data held within this warehouse 'research-ready'. The PATHS Data Resource was assembled from Repository data to facilitate child health research and is a distinct subset of the Repository; as such, investigators using the PATHS Resource will have minimal upfront work to get the data 'research-ready'.

PATHS includes over 99% of all children residing in Manitoba, encompassing all individuals registered with Manitoba's universal health insurance system at some point during childhood (0–18 years), born between 1984 and 2012. Child data begin at birth (e.g. birthweight, gestational age, 5-min Apgar score, socioeconomic status at birth) and continue until (i) the child dies or moves out of the province, or (ii) the latest data acquisition in 2012. Data on maternal health during the preconception period through to delivery are also available.

### PATHS: A unique opportunity

The PATHS Resource offers a unique opportunity to study child health and health equity: (i) it is unique in the breadth of measures it contains; (ii) it follows children from the prenatal period; (iii) the data are linkable across domains (health, social services, education); (iv) the data are

linkable within families, providing the ability to identify siblings.

### PATHS Data Resource: funding sources

For over 20 years, the Manitoba government has funded MCHP to maintain the Repository; the Repository currently comprises over 70 databases from health, education, social services and justice, as well as registry data. Funds from the Canada Foundation for Innovation have supported efforts to bring in additional databases over the past 4 years [Grant Number CFI20329]. In 2011, grants from the Canadian Institutes for Health Research [CIHR:ROH-115206] and the Heart and Stroke Foundation [PG-12-0534] enabled the creation of the PATHS Data Resource, drawing data held in the Repository from a variety of sectors.

## Data Resource Area and Population Coverage

### Identifying children to include in the PATHS Data Resource

Children were identified for PATHS using the Manitoba Health Insurance Registry, a database of all individuals registered with Manitoba's universal health insurance programme. All individuals with a birth date between 1 April 1984 and 31 March 2012 were extracted for inclusion in PATHS, regardless of place of birth ( $n=615\,128$ ). We excluded individuals who did not reside in Manitoba at any point during childhood age 0–18 years ( $n=30\,873$ ). The final PATHS Resource comprises data on 584 255 children.

Tables 1 and 2 provide a flavour of the capabilities of the PATHS Resource. Table 1 presents summary characteristics for selected birth cohorts, 1984–2012. The median maternal age at first birth ranged from a high of 26.5 years (1984/85 to 1985/86) to a low of 23.8 years (2007/08 to 2008/09). Table 2 presents examples of time-varying characteristics for each birth cohort. The data show that children are more likely to live in households receiving income assistance during their early childhood (ages 0–6 years) compared with when they are older. Numerous other measures are available within PATHS.

### Coverage

The PATHS data are population-based, representing 99+ % of individuals who lived in Manitoba at some point during childhood. These data allow researchers to track

marginalized families that may otherwise be lost to follow-up in other research approaches or surveys.<sup>10,17,18</sup>

## Measures

The PATHS Resource includes data from several sectors, that were routinely collected on the population of Manitoban children for administrative purposes; they provide a cost-effective opportunity for conducting epidemiological research.<sup>17–25</sup>

### Sources of data

The data are provided by various entities (e.g. ministries of the Government of Manitoba). Data providers are owners of the data held within the PATHS Resource; MCHP is the steward of these data. Figure 2 illustrates the cross-cutting, longitudinal structure of the PATHS Resource.

**Manitoba Health Registry.** This data source includes virtually every single child residing in Manitoba; i.e. all children registered with Manitoba's universal health insurance system. It contains data on age, sex, family composition, date of birth, postal code of residence, death and migration in/out of province.

**Hospital Discharge Abstract Database (HDAD).** The HDAD comprises data from hospital records containing summaries of demographic and clinical information that were abstracted at hospital discharge: diagnostic codes (International Classification of Disease), procedure codes (e.g. Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures, and Canadian Classification of Health Interventions), sex and postal code. The HDAD includes data for all Manitobans discharged from acute and chronic care facilities both in and outside Manitoba.

**Physician claims.** These data include claims for physician office visits, fee-for-service components for tests (e.g. lab and X-ray procedures), and three-digit ICD diagnosis associated with the physician visit.

**Drug Program Information Network (DPIN).** DPIN is a point-of-sale prescription drug database. It contains complete drug profiles for each client for all prescriptions in Manitoba, regardless of insurance coverage. DPIN does not capture prescriptions provided in hospital or at nursing stations, nor outpatient visits at CancerCare Manitoba.

**Manitoba Immunization Monitoring System (MIMS).** MIMS data include information on the type of vaccine, service date and information about the provider administering the immunization.

**Healthy Child Manitoba Office (HCMO).** HCMO provides a variety of policies and programmes to promote optimal health outcomes in children. Data from HCMO include biological, social and demographic risk factors for

**Table 1.** Selected birth characteristics of children from five birth cohorts within the PATHS data resource

Child characteristics at birth	Selected birth cohorts (by fiscal year)					
	1984/85 to 1985/86 <i>n</i> = 33 085 (Std Error)	1992/93 to 1993/94 <i>n</i> = 32 896 (Std Error)	1997/98 to 1998/99 <i>n</i> = 28 861 (Std Error)	2002/03 to 2003/04 <i>n</i> = 27 696 (Std Error)	2007/08 to 2008/09 <i>n</i> = 30 730 (Std Error)	2010/11 to 2011/12 <i>n</i> = 31 455 (Std Error)
Maternal age at first birth, median years	26.5 (0.03)	25.8 (0.03)	24.3 (0.03)	24.0 (0.03)	23.8 (0.03)	24.3 (0.03)
Income quintile at birth						
Income quintile 5, percentage	18.3 (0.21)	17.1 (0.20)	16.0 (0.22)	16.0 (0.22)	15.9 (0.21)	7.6 (0.15)
Income quintile 4, percentage	20.3 (0.22)	19.5 (0.22)	18.2 (0.23)	18.0 (0.23)	16.8 (0.21)	8.7 (0.16)
Income quintile 3, percentage	18.5 (0.21)	19.1 (0.21)	18.6 (0.23)	18.8 (0.23)	18.9 (0.22)	9.4 (0.16)
Income quintile 2, percentage	19.4 (0.22)	19.7 (0.22)	21.1 (0.24)	20.2 (0.24)	20.9 (0.23)	10.5 (0.17)
Income quintile 1, percentage	23.2 (0.23)	24.0 (0.23)	25.6 (0.26)	26.7 (0.27)	27.2 (0.25)	13.3 (0.19)
Income quintile missing, percentage	0.3 (0.03)	0.6 (0.04)	0.4 (0.04)	0.3 (0.03)	0.3 (0.03)	0.2 (0.02)
Gestational age at birth						
Gestation age <37 weeks, percentage	6.0 (0.13)	6.3 (0.13)	6.9 (0.15)	7.3 (0.16)	7.6 (0.15)	7.3 (0.15)
Gestation age ≥37 weeks, percentage	91.7 (0.15)	91.8 (0.15)	90.7 (0.17)	89.7 (0.18)	89.0 (0.18)	88.4 (0.18)
Gestation missing, percentage	2.2 (0.08)	2.0 (0.08)	2.4 (0.09)	3.0 (0.10)	3.3 (0.10)	4.3 (0.11)
5 minute Apgar score (0–10)						
Apgar ≤8, percentage	15.0 (0.20)	15.4 (0.20)	14.7 (0.21)	13.6 (0.21)	12.3 (0.19)	12.5 (0.19)
Apgar >8, percentage	82.7 (0.21)	82.6 (0.21)	82.8 (0.22)	83.3 (0.22)	84.3 (0.21)	83.1 (0.21)
Apgar missing, percentage	2.3 (0.08)	2.0 (0.08)	2.6 (0.09)	3.1 (0.10)	3.5 (0.10)	4.3 (0.11)
Birthweight						
Birthweight <3500 g, percentage	51.6 (0.27)	52.1 (0.27)	50.2 (0.29)	48.4 (0.30)	50.2 (0.28)	50.9 (0.28)
Birthweight ≥3500 g, percentage	46.3 (0.27)	46.0 (0.27)	47.3 (0.29)	48.7 (0.30)	46.9 (0.28)	45.3 (0.28)
Birthweight missing, percentage	2.1 (0.08)	1.9 (0.07)	2.5 (0.09)	3.0 (0.10)	2.8 (0.09)	3.8 (0.11)
Birth order of child						
Child is firstborn, percentage	33.3 (0.26)	32.1 (0.25)	29.9 (0.27)	30.0 (0.27)	28.6 (0.26)	28.6 (0.25)
Child is second born, percentage	32.4 (0.26)	31.6 (0.25)	29.7 (0.27)	28.9 (0.27)	27.3 (0.25)	26.8 (0.25)
Child is third born or later, percentage	33.3 (0.26)	35.4 (0.26)	39.1 (0.29)	40.0 (0.29)	40.8 (0.28)	40.8 (0.28)
Birth order missing, percentage	1.0 (0.05)	0.9 (0.05)	1.3 (0.07)	1.1 (0.06)	3.3 (0.10)	3.8 (0.11)

Std Error, standard error.

**Table 2.** Select time-varying characteristics for seven birth cohorts in PATHS at each of 9 fiscal years

Characteristics	1995/96	1997/98	1999/2000	2001/02	2003/04	2005/06	2007/08	2009/10	2011/12							
	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)							
Number of hospitalizations for injury																
Birth cohorts																
1984/85 to 1985/86 ( <i>n</i> = 33 375)	159	215	265	309	304	328	283	277	298							
1987/88 to 1988/89 ( <i>n</i> = 33 773)	149	133	123	190	227	363	349	335	325							
1992/93 to 1993/94 ( <i>n</i> = 33 134)	210	173	163	100	112	161	267	296	293							
1997/98 to 1998/99 ( <i>n</i> = 28 967)	61	61	144	123	120	94	125	105	163							
2002/03 to 2003/04 ( <i>n</i> = 27 774)					106	144	100	94	117							
2007/08 to 2008/09 ( <i>n</i> = 30 850)							82	187	139							
2010/11 to 2011/12 ( <i>n</i> = 31 585)									178							
Median number of physician visits																
Birth cohorts																
1984/85 to 1985/86 ( <i>n</i> = 33 375)	2	(0.11)	1	(0.14)	2	(0.11)	1	(0.10)	2	(0.07)	2	(0.06)	3	(0.04)		
1987/88 to 1988/89 ( <i>n</i> = 33 773)	2	(0.11)	2	(0.13)	1	(0.22)	1	(0.16)	1	(0.08)	1	(0.06)	1	(0.04)	3	(0.04)
1992/93 to 1993/94 ( <i>n</i> = 33 134)	2	(0.11)	2	(0.14)	1	(0.19)	1	(0.08)	1	(0.07)	1	(0.07)	1	(0.04)	3	(0.03)
1997/98 to 1998/99 ( <i>n</i> = 28 967)	4	(0.22)	2	(0.15)	2	(0.10)	2	(0.10)	2	(0.07)	1	(0.04)	1	(0.03)	2	(0.02)
2002/03 to 2003/04 ( <i>n</i> = 27 774)					3	(0.18)	2	(0.09)	1	(0.05)	1	(0.03)	1	(0.03)	2	(0.02)
2007/08 to 2008/09 ( <i>n</i> = 30 850)							2	(0.12)	2	(0.05)	2	(0.05)	2	(0.05)	3	(0.02)
2010/11 to 2011/12 ( <i>n</i> = 31 585)															5	(0.02)

(Continued)

**Table 2. Continued**

Characteristics	1995/96	1997/98	1999/2000	2001/02	2003/04	2005/06	2007/08	2009/10	2011/12
	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)
Percent of children with ADHD diagnosis									
Birth cohorts									
1984/85 to 1985/86 (n = 33 375)	1.3% (0.06)	1.7% (0.07)	1.5% (0.07)	0.9% (0.05)	0.5% (0.04)	0.3% (0.03)	0.3% (0.03)	0.3% (0.03)	0.4% (0.03)
1987/88 to 1988/89 (n = 33 773)	1.0% (0.05)	2.1% (0.08)	2.4% (0.08)	2.2% (0.08)	1.8% (0.07)	1.1% (0.06)	0.5% (0.04)	0.5% (0.04)	0.6% (0.04)
1992/93 to 1993/94 (n = 33 134)	0.0% (0.00)	0.1% (0.02)	0.8% (0.05)	2.1% (0.08)	3.0% (0.09)	3.1% (0.09)	2.7% (0.09)	2.0% (0.08)	1.4% (0.07)
1997/98 to 1998/99 (n = 28 967)				0.0% (0.01)	0.7% (0.05)	2.3% (0.09)	3.3% (0.11)	3.9% (0.11)	4.0% (0.12)
2002/03 to 2003/04 (n = 27 774)					0.0% (0.00)	0.0% (0.01)	0.3% (0.03)	1.7% (0.08)	3.9% (0.12)
2007/08 to 2008/09 (n = 30 850)								0.0% (0.01)	0.1% (0.02)
2010/11 to 2011/12 (n = 31 585)									
Percent of children living in households receiving income assistance									
Birth cohorts									
1984/85 to 1985/86 (n = 33 375)	8.0% (0.15)	6.8% (0.14)	6.5% (0.13)	3.3% (0.10)					
1987/88 to 1988/89 (n = 33 773)	10.0% (0.16)	8.4% (0.15)	7.9% (0.15)	6.5% (0.13)	5.8% (0.13)				
1992/93 to 1993/94 (n = 33 134)	14.8% (0.20)	12.6% (0.18)	11.8% (0.18)	9.6% (0.16)	8.5% (0.15)	7.9% (0.15)	6.6% (0.14)	3.8% (0.10)	
1997/98 to 1998/99 (n = 28 967)		5.1% (0.13)	16.5% (0.22)	14.2% (0.21)	12.4% (0.19)	11.2% (0.19)	9.3% (0.17)	8.0% (0.16)	7.7% (0.16)
2002/03 to 2003/04 (n = 27 774)					13.3% (0.20)	15.5% (0.22)	12.8% (0.20)	11.2% (0.19)	10.3% (0.18)
2007/08 to 2008/09 (n = 30 850)							4.8% (0.12)	14.4% (0.20)	13.6% (0.19)
2010/11 to 2011/12 (n = 31 585)									11.4% (0.18)

(Continued)

Table 2. Continued

Characteristics	1995/96	1997/98	1999/2000	2001/02	2003/04	2005/06	2007/08	2009/10	2011/12
	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)	(Std Error)
Percentage of households moved more than once in a given year									
Birth cohorts									
1984/85 to 1985/86 ( <i>n</i> = 33 375)	0.6% (0.04)	0.4% (0.04)	0.4% (0.03)	0.4% (0.03)	0.5% (0.04)	0.8% (0.05)	0.6% (0.04)	0.6% (0.04)	0.6% (0.04)
1987/88 to 1988/89 ( <i>n</i> = 33 773)	0.9% (0.05)	0.6% (0.04)	0.5% (0.04)	0.5% (0.04)	0.3% (0.03)	0.5% (0.04)	0.7% (0.04)	0.6% (0.04)	0.6% (0.04)
1992/93 to 1993/94 ( <i>n</i> = 33 134)	1.8% (0.07)	1.0% (0.05)	0.8% (0.05)	0.6% (0.04)	0.5% (0.04)	0.6% (0.04)	0.5% (0.04)	0.4% (0.04)	0.4% (0.04)
1997/98 to 1998/99 ( <i>n</i> = 28 967)		1.0% (0.06)	1.1% (0.06)	1.1% (0.06)	0.8% (0.05)	0.9% (0.06)	0.6% (0.05)	0.6% (0.05)	0.6% (0.05)
2002/03 to 2003/04 ( <i>n</i> = 27 774)					0.9% (0.06)	1.3% (0.07)	1.1% (0.06)	0.8% (0.05)	0.8% (0.05)
2007/08 to 2008/09 ( <i>n</i> = 30 850)								1.0% (0.06)	1.0% (0.06)
2010/11 to 2011/12 ( <i>n</i> = 31 585) <sup>a</sup>									

ADHD, attention deficit-hyperactivity disorder.

<sup>a</sup>Data on residential mobility, 2010/11 to 2011/12, not available in most recent data update. Data will be included in next update 2014.



**Figure 1.** Province of Manitoba: the PATHS Data Resource includes population-based data on nearly every child born in Manitoba, 1984–2012<sup>40</sup>.

children such as prenatal exposures (e.g. alcohol, tobacco), parental anxiety disorders, parental education and criminal involvement. HCMO data also include the Early Development Instrument which assesses children's developmental status at kindergarten across five areas: (i) physical health, (ii) communication skills and knowledge, (iii) social competence, (iv) emotional maturity and (v) language and cognitive development.

**Tenant management system.** This database provides details on all households residing in units managed by Manitoba Housing, which offers affordable housing to vulnerable families. Data include demographic information (age, sex, marital status), type of household, length of stay in Manitoba Housing, building type and location of building.

**Child and family services.** Data from child and family services contain information on children receiving child welfare services from Manitoba Family Services. Measures include out-of-home care, protection orders and measures of maltreatment allegations and substantiations.

**Education.** Education data provide information on enrolment, special needs funding, courses, marks, standards tests, graduation and demographics. Data are available from both public and private schools, as well as those who are home schooled.

**Jobs and the Economy.** Data from Manitoba Jobs and the Economy contain information on families receiving employment and income assistance (i.e. welfare).

**Canadian Census Data.** Income data come from the Canadian Census. The Census, conducted by Statistics Canada, includes all Canadian citizens, landed immigrants and non-permanent residents.<sup>26</sup> Measures on income are

based on average household income (as reported on the Census) for each census dissemination area, each of which covers approximately 400–700 individuals. Average income for census dissemination areas has been validated as a measure of socioeconomic status.<sup>27</sup>

### Data linkage

**Figure 3** summarizes the data linkage process. All persons registered with Manitoba's universal health insurance programme (Manitoba Health) receive a nine-digit Personal Health Identification Number (PHIN). Health Information Management (HIM) uses each registrant's PHIN to create a unique, encrypted PHIN for the Repository at MCHP. HIM replaces each individual's actual PHIN with their encrypted PHIN and deletes identifying information (e.g. name, address, actual PHIN) from the record. The de-identified data are then sent to a secure data environment within MCHP.

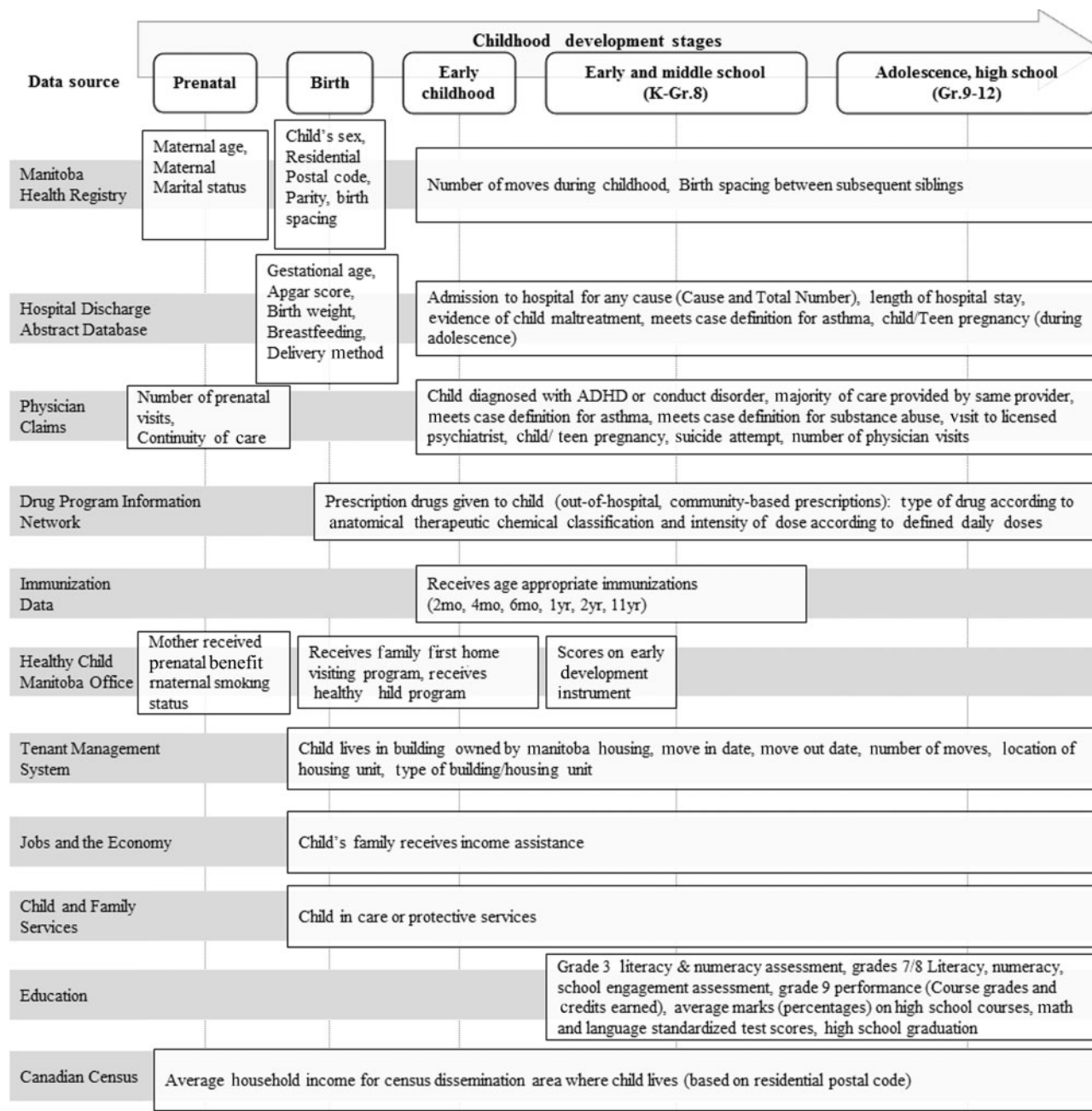
Data from non-health ministries (e.g. education, social services etc.) are linked using a different process. Ministries first generate two individual-level data files. The first includes individuals' identifying information (e.g. name, birth date) and each individual's encrypted programme number; no programme data is included. The second data file contains the encrypted programme number and all programmatic data, with no identifying information. The first data file is sent to HIM where identifiers (e.g. name, birth date) are used to link individuals to their encrypted PHIN. HIM next strips all identifying information, leaving only the encrypted PHINs and programme numbers in a crosswalk file. The second data set is sent to MCHP, who use this crosswalk file to attach the encrypted PHIN to the programme data, and to link these non-health data (e.g. education data) with health data.

Data are linkable at the individual level across all databases within PATHS. Linkages can only be performed on approved projects, within the secure data environment at MCHP or at remote access sites (currently six are located in Manitoba, Canada).

### MCHP Metadata

Linking data at the individual level is only part of the process for transforming PATHS data into research data. The other aspect involves interacting with and applying the MCHP Metadata—a comprehensive set of documentation and tools that describes each database, data quality reports, measures included in the databases, SAS macros<sup>28</sup> for generating variables, data summaries and timelines, and more. The Metadata allow the PATHS data to be consistently used across various epidemiological studies.





**Figure 2.** Longitudinal nature of the multiple linked data sources in the PATHS Data Resource. Examples of measures for each data source over time are presented. A complete listing of all measures available in the PATHS data resource is available from the authors.

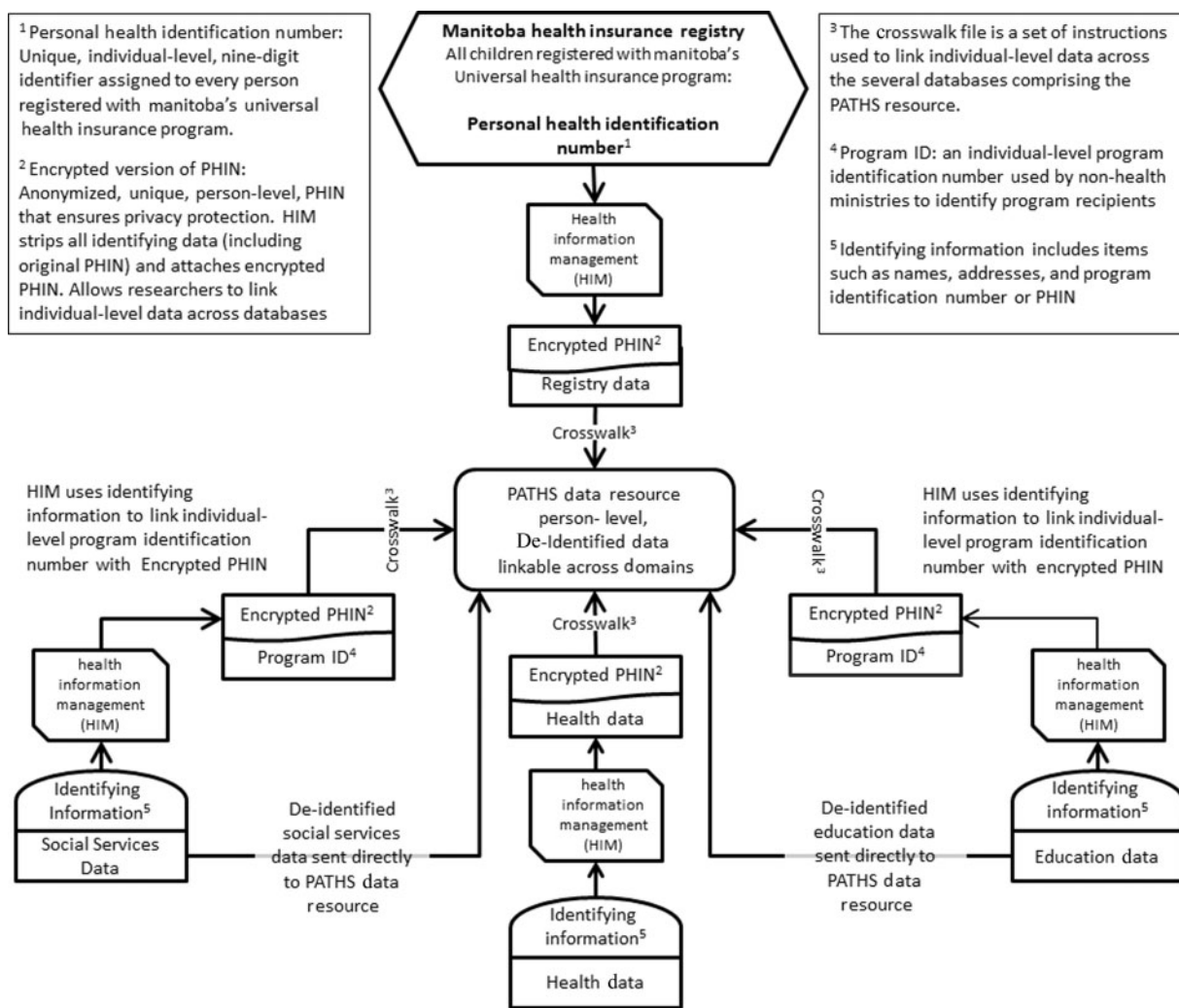
One of the Metadata resources central to this process is the MCHP Concept Dictionary.<sup>29,30</sup> This resource provides researchers with information about (i) measures included in the PATHS Resource, (ii) validated algorithms for administrative data, (iii) the operationalization of constructs and variables, (iv) cautions and recommendations regarding SAS<sup>28</sup> programming and (v) references for each concept.

Another key feature of the MCHP Metadata is that researchers and analysts are able to provide feedback and ask questions via an interactive interface, facilitating

real-time information flow between researchers, analysts and data management experts.

### Frequency of Data Collection/Update

Survey data often rely on intermittent contact with study participants. The PATHS Resource, however, contains data that are continuously collected; i.e. whenever a child comes into contact with the healthcare system, whenever the child's school marks are recorded and/or whenever the



**Figure 3.** Data elements used to link individual-level data from multiple databases across several domains.

child comes into contact with social services. Data are sent to HIM for (i) de-identification and (ii) attachment of encrypted PHIN on an annual basis.

The median observation time for a child included in PATHS, from birth, is 15.4 years (inter-quartile range 8.1 to 21.8 years).

**Data Resource Use: Using the Paths Data**

The PATHS Resource is currently being used to evaluate 12 interventions, programmes or policies aimed at improving health or educational outcomes for children. In all cases, it was used to identify the children exposed to the intervention or programme, determine which relevant outcomes were available and create a flat file for analysis. The interventions include both universal and targeted programmes, some aimed at specific age groups or specific health disorders or a particular portion of the socio-economic stratum. Current and completed analyses include:

evaluating the impact of the location of social housing on the health and educational outcomes of residents; a pre-natal benefit for low-income pregnant women on birth outcomes and infant health; and also, evaluating the impact of a pay-for-performance physician-funding model on immunization and vaccination rates. At the time of writing, one article had been published using these data<sup>31</sup> and others are currently under review or have been recently accepted.<sup>32,33</sup>

**Strengths and Weaknesses**

The PATHS Resource has limitations. Not all important health data are captured in administrative data (e.g. body mass index). For individuals who have left the province or moved into the province sometime after birth, any relevant contacts while out of the province are permanently missing. Individual data sets may also have some particular weaknesses. For example, medical claims data (i.e.

physician visits) only include a single three-digit ICD9 diagnosis and may miss some comorbid conditions.

Despite some limitations, the PATHS Resource has several strengths. First and foremost, the data are universal, in that the entire population of Manitoba is captured and present. Second, the PATHS Resource is longitudinal, with virtually complete capture of contacts with the healthcare system including prescription drug dispensations, and data from other domains, such as school enrolment or receipt of social assistance. Third, it is linkable across domains and over time. The data have also been validated in several studies.<sup>18,19,34–39</sup>

### Accessing the Paths Data Resource

MCHP welcomes researchers who study child health and development to apply for access to the PATHS Resource. There are four requirements for accessing data held within the Repository, which includes the PATHS Resource. The first requirement is that all researchers using the PATHS Resource complete the MCHP Accreditation which provides a consistent overview of MCHP, data access and available data. The second is a feasibility assessment conducted by MCHP, which ensures that the data elements required for the research question are present in the data and that the proposed project fits within the MCHP mandate. Third, the investigators must have their project approved by (i) the university research ethics board, (ii) the government-mandated Health Information Privacy Committee and, where appropriate, (iii) non-health data providers, to ensure the proper handling of the data and reporting of research results. Fourth, after all approvals have been obtained, the investigators will complete a researcher agreement between the University of Manitoba and Manitoba Health, Healthy Living & Seniors. On a strictly cost recovery basis, there are some charges related to the administration and housing of the data at MCHP.

Investigators may travel to one of the remote access sites to work directly with the data. Alternatively, investigators may work with an MCHP-based analyst to carry out the proposed study. Investigators can coordinate with the analyst to develop analytical procedures to address the proposed research objectives. Once the procedures are specified, the analyst will carry out analyses on the linked, individual-level data comprising the PATHS Resource within the secure data environment at MCHP. The analyst will then share the procedures and results according to the research agreement mentioned above. The sensitivity and breadth of the PATHS data, in terms of individuals covered and measures included, require that the data remain within the secure MCHP environment, as per policies and procedures in place. Moreover, the data-sharing agreements

between MCHP and data providers (i.e. owners of the data) state that individual-level data and data summaries based on counts fewer than five cannot be sent outside MCHP. That said, hundreds of projects have been conducted by external investigators (i.e. non MCHP faculty) from around the globe using data held within the Repository, of which the PATHS Resource is a distinct subset. The experiences of these investigators demonstrate that using data held in the Repository, including the PATHS Resource, is not onerous to interested researchers. Researchers interested in using the PATHS Resource for their own research are encouraged to visit [[http://www.umanitoba.ca/faculties/medicine/units/community\\_health\\_sciences/departamental\\_units/mchp/protocol/media/PATHS\\_Resource\\_Website.pdf](http://www.umanitoba.ca/faculties/medicine/units/community_health_sciences/departamental_units/mchp/protocol/media/PATHS_Resource_Website.pdf)] and/or e-mail [mchp\_access@cpe.umanitoba.ca] to learn how to apply for access.

### Supplementary Data

Supplementary data are available at *IJE* online.

### Funding

P.M. wishes to acknowledge funding from the Canadian Institutes of Health Research (CIHR) and the Public Health Agency of Canada (PHAC) for her CIHR/PHAC Applied Public Health Research Chair (2008–2014). M.B. acknowledges the financial support of the Government of Manitoba through the Manitoba Center for Health Policy Population-Based Child Health Research Award.

### Acknowledgements

This work was supported by the Canadian Institutes of Health Research (CIHR) and the Heart & Stroke Foundation of Canada, under the programme of research entitled 'PATHS Equity for Children: a program of research into what works to reduce the gap for Manitoba's children'. The authors acknowledge the Manitoba Centre for Health Policy (MCHP) for use of data contained in the Population Health Research Data Repository under project # 2011-027 (HIPC #2011/2012 – 24). The results and conclusions are those of the authors and no official endorsement by MCHP, Manitoba Health, Healthy Living and Seniors (MHLS) or other data providers is intended or should be inferred. The PATHS Equity team includes: James Bolton; Marni Brownell; Charles Burchill; Elaine Burland; Mariette Chartier; Dan Chateau; Malcolm Doupe; Greg Finlayson; Randall Fransoo; Chun Yan Goh; Mingming Hu; Doug Jutte; Alan Katz; Laurence Katz; Lisa Lix; Patricia J Martens; Colleen Metge; Nathan C Nickel; Colette Raymond; Les Roos; Noralou Roos; Rob Santos; Joykrishna Sarkar; Mark Smith; Carole Taylor; Randy Walld.

**Conflict of interest:** None declared.

### References

1. Syme SL, Berkman LF. Social class, susceptibility and sickness. *Am J Epidemiol* 1976;**104**:1–8.
2. Marmot MG, Stansfeld S, Patel C *et al.* Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991;**337**:1387–93.

3. Marmot MG, Kogevinas M, Elston MA. Social/economic status and disease. *Annu Rev Public Health* 1987;8:111–35.
4. Adler NE, Boyce WT, Chesney MA, Folkman S, Syme SL. Socioeconomic inequalities in health: No easy solution. *JAMA* 1993;26:3140–45.
5. Tudor Hart J. The Inverse Care Law. *Lancet* 1971;297:405–12.
6. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet* 2008;372:1661–69.
7. Marmot M, Bell R. Fair society, healthy lives. *Public Health* 2012;126(Suppl 1):S4–10.
8. Chumney EC, Mauldin PD, Simpson KN. Charges for hospital admissions attributable to health disparities for African-American patients, 1998–2002. *J Natl Med Assoc* 2006;98:690–94.
9. Mackenbach JP, Meerding WJ, Kunst AE. Economic costs of health inequalities in the European Union. *J Epidemiol Community Health* 2011;65:412–19.
10. Roos NP, Sullivan K, Walld R, MacWilliam L. Potential savings from reducing inequalities in health. *Can J Public Health* 2004;95:460–64.
11. Schoeni RF, Dow WH, Miller WD, Pamuk ER. The economic value of improving the health of disadvantaged Americans. *Am J Prev Med* 2011;40(Suppl 1):S67–S72.
12. Mills C, Reid P, Vaithianathan R. The cost of child health inequalities in Aotearoa New Zealand: a preliminary scoping study. *BMC Public Health* 2012;12:384.
13. Victora CG, Wagstaff A, Schellenberg JA, Gwatkin D, Claeson M, Habicht JP. Applying an equity lens to child health and mortality: more of the same is not enough. *Lancet* 2003;362:233–41.
14. Hertzman C. Population health and human development. In: Keating DP, Hertzman C (eds). *Developmental Health and the Wealth of Nations*. New York: Guilford Press, 1999.
15. Starfield B. Pathways of influence on equity in health. *Social Science & Medicine* 2007;64:1355–62.
16. Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, Molecular Biology, and the Childhood Roots of Health Disparities Building a New Framework for Health Promotion and Disease Prevention. *Jama-Journal of the American Medical Association*. 2009 Jun 3;301:2252–9.
17. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health* 2011;32:91–108.
18. Roos LL, Nicol JP. A research registry: Uses, development, and accuracy. *J Clin Epidemiol* 1999;52:39–47.
19. Roos LL, Brownell M, Lix L, Roos NP, Walld R, MacWilliam L. From health research to social research: Privacy, methods, approaches. *Soc Sci Med* 2008;66:117–29.
20. Roos NP, Roos LL, Brownell M, Fuller EL. Enhancing policy-makers' understanding of disparities: relevant data from an information-rich environment. *Milbank Q* 2010;88:382–403.
21. Shu X, Ji J, Li X, Sundquist J, Sundquist K, Hemminki K. Cancer risk in patients hospitalised for Graves' disease: a population-based cohort study in Sweden. *Br J Cancer* 2010;102:1397–99.
22. Black SE, Devereux PJ, Salvanes KG. From the cradle to the labor market? The effect of birthweight on adult outcomes. *Q J Econ* 2007;122:409–39.
23. Goldacre MJ, Griffith M, Gill L, Mackintosh A. In-hospital deaths as fraction of all deaths within 30 days of hospital admission for surgery: analysis of routine statistics. *BMJ* 2002;324:1069–70.
24. Goldacre M, Kurina L, Yeates D, Seagroatt V, Gill L. Use of large medical databases to study associations between diseases. *QJM* 2000;93:669–75.
25. Jutte DP, Brownell M, Roos NP, Schippers C, Boyce WT, Syme SL. Rethinking what is important: biologic versus social predictors of childhood health and educational outcomes. *Epidemiology* 2010;21:314–23.
26. Statistics Canada. *Census of Population 2011*. Ottawa ON: Statistics Canada, 2011.
27. Chateau D, Metge C, Prior H, Soodeen RA. Learning from the census: the Socio-economic Factor Index (SEFI) and health outcomes in Manitoba. *Can J Public Health* 2012;103(Suppl 2):S23–S27.
28. *Version 9.2 of the SAS System for Windows*. Cary, NC: SAS Institute, 2008.
29. Burchill C, Roos LL, Fergusson P, Jebamani L, Turner K, Dueck S. Organizing the present, looking to the future: an online knowledge repository to facilitate collaboration. *J Med Internet Res* 2000;2:E10.
30. Roos LL, Soodeen RA, Bond R, Burchill C. Working more productively: Tools for administrative data. *Health Serv Res* 2003;38:1339–57.
31. Brownell MD, Nickel NC, Chateau D *et al*. Long-term benefits of full-day kindergarten: a longitudinal population-based study. *Early Child Dev Care* 2014; doi: 10.1080/03004430.2014.913586.
32. Nickel NC, Martens PJ, Chateau D *et al*. Have we left some behind? Trends in socioeconomic inequalities in breastfeeding initiation: A population-based epidemiological surveillance study. *Can J Public Health* 2014. In press.
33. Martens PJ, Chateau D, Burland E *et al*. The effect of neighborhood socioeconomic status on education and health outcomes for children living in social housing. *Am J Public Health* 2014. In press. doi: 10.2105/AJPH.2014.302133.
34. Roos LL, Gupta S, Soodeen RA, Jebamani L. Data quality in an information-rich environment: Canada as an example. *Can J Aging* 2005;24(Suppl 1):153–70.
35. Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease. Comparing administrative data and self-reports. *Med Care* 1997;35:932–47.
36. Oreopoulos P, Stabile M, Walld R, Roos LL. Short-, medium-, and long-term consequences of poor infant health. An analysis using siblings and twins. *J Hum Resources* 2008; 43:88–138.
37. Kozyrskyj AL, Mustard CA. Validation of an electronic, population-based prescription database. *Ann Pharmacother* 1998;32:1152–57.
38. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health* 2011;32:91–108.
39. Brownell MD, Roos NP, Roos LL. Monitoring health reform: a report card approach. *Soc Sci Med* 2001;52:657–70.
40. Statistics Canada. *Canada 2006*. Winnipeg, MB: Manitoba Centre for Health Policy, University of Manitoba, 2006.