

---

## Inventory of National Rural Health Research Related Databases

---

**Database name:** **National Longitudinal Survey of Children and Youth**

**Thematic Coverage:** This survey follows a representative sample of Canadian children, aged newborn to 11 years, into adulthood, with data collection occurring at two-year intervals. It is a single source of data for the examination of child development in context, including the diverse life paths of normal development. The primary objective is to inform policy by establishing a national database on the characteristics and life experiences of children and youth in Canada as they grow up.

**Maintained by:** Special Surveys Division, Statistics Canada

**Availability:** A Public version of the file is available through the Data Liberation Initiative (DLI). Some variables are suppressed and other are aggregated to protect the anonymity of individual survey respondents. Custom tabulations are available for the master file on a cost recovery basis

---

**Start Date:** 1994  
**Release Date:** Approximately 12 months after data collection  
**Frequency of collection:** Biannual

---

**Data Collection:** Using the Labour Force Survey as a sampling frame, a parent of the child asked to complete questionnaires for children 10 years and older. A mail follow-up with the teacher and the school principal and an achievement test was also collected.

**Sample size:** Approximately 15,000; Younger children are being added to the sample as the initial cohort ages; target sample size = 25,000 children.

For the NLSCY, the sample was constructed taking account the following requirements.

- A sufficient sample was required in each of the 10 provinces to allow for the production of reliable estimates for all longitudinal children who were 0 to 11 years of age in Cycle 1.
- It was also necessary to have a large enough sample to produce estimates for Cycle 1 at the Canada level by seven key age groupings or cohorts: 0 to 11 months, 1 year, 2 to 3 years, 4 to 5 years, 6 to 7 years, 8 to 9 years, and 10 to 11 years.
- In each province, a sufficient sample size was required for Cycle 2 to produce reliable estimates for all children who were 0 to 11 years of age in Cycle 1.

---

**Geographic coverage:** All provinces and territories, excluding Aboriginal children living on Reserves and children in institutions.

**Lowest geographic level collected:** Postal Code

Lowest geographic level of release: Province/Territory. Data is suppressed for unweighted cells with less than 30 observations. While the NLSCY national sample size is a strength, provincial sample sizes are too small to conduct analyses for specific age groups at the provincial level.

Existing rural variable: No

Rural definitions that can be constructed from this database include (building block)<sup>†</sup>:

- ✓ Census "rural areas" (Enumeration Area)
- ✓ Rural and Small Town definition (Census Subdivision)
- ✓ Metropolitan area and census agglomeration Influenced Zones (Census Subdivision)
- ✓ OECD "rural communities" definition (Census Consolidated Subdivision)
- ✓ OECD "predominantly rural regions" definition (Census Division)
- ✓ Ehrensaft's "Beale codes" (Census Division)

<sup>†</sup> Results for these areas of geography could conceivably be derived from postal codes if respondent confidentiality is ensured

---

Data Elements:

- Outcomes (e.g. health, language, cognitive, social, emotional, behavioural)
- Characteristics of child's family (SES, structure, parenting style, family functioning, social support)
- Child care, School, Neighbourhood
- Direct assessments of child (e.g. tests of vocabulary, math and reading skills)
- Reports from parents, teachers, principals

Notes:

Use of the Labour Force Survey sampling frame does not capture the children who are most at risk such as those in homeless or transient families and makes it difficult for the NLSCY to over-sample certain key populations, such as infants.

**Warning:** Regarding the nature of the primary and self files in **Cycles 2 and 3** of the **NLSCY**. Unlike Cycle 1, where the primary and secondary files can be linked through the use of CHILDDID, the records in the files for Cycles 2 and 3 **cannot** be merged. It was assumed that the CHILDDID variable in these two files could be used to match cases between them. This is a natural enough assumption, since both files in Cycle 1 can be linked using CHILDDID. It therefore would seem reasonable to assume that both files in Cycle 3, which also have CHILDDID, could be merged. **This is not the case.** This misunderstanding is compounded by the fact that the documentation for Cycle 3 does not state anywhere that the primary and 10-15 files cannot be linked.

The user's guide for Cycle 2, however, contains the following warning.

*"The data for the self-completed portion of the file was separated from the main file to ensure that the information provided by children was kept confidential, even from their parents. In order to protect that confidentiality, the information from both files cannot be linked. An assessment and further suppressions were made following the procedures outlined in the next pages to ensure that the two files could not be linked."* p. 141