

Child and Youth Development Research Partnership

Key messages

The known:

- Data linkage is a method for bringing together existing information about the same person from multiple data sources. It involves strict principles of protecting privacy by the separation of “identifying” information from “content” information
- Data linkage provides answers to research questions which are based on large-scale studies of a whole population and not on surveys or samples.

The new:

- In the Northern Territory, Menzies School of Health Research and seven NT Government departments have collaborated to develop an extensive, de-identified “data repository” containing information on children and young people to be used for research to guide policy and services to improve outcomes for young Territorians.

The implications:

- Researchers are using the data repository for a wide range of projects, with aims that include:
 - reducing premature birth
 - supporting early child development
 - improving school attendance
 - understanding factors that influence the mental health of young people
 - reducing youth offending, and
 - understanding the influences on life course trajectories for young Territorians

What is data linkage?

Many of our life experiences, from birth through to death are recorded as “data” in a wide range of administrative datasets. For example, information is collected when children are born, when they attend school or when they visit a health clinic. Data linkage is a method for bringing together, or “linking”, existing information about the same person from multiple sources. Bringing together information for many individuals allows researchers to understand important factors which may influence results across a wide range of social outcomes. Importantly there are strict protocols to ensure individual privacy is protected.

Linking data can be used for a wide range of research:

- investigating the health and wellbeing of a population
- evaluation of policies, services or the impact of major events
- studying factors associated with health and social outcomes
- understanding the multiple influences on life course trajectories.

What is CYDRP ?

In 2009, a group of government agencies and universities in South Australia and the Northern Territory (NT) agreed to build a resource to support research using linked data, now known as SA NT DataLink. SA NT DataLink does not do research but works with data custodians and researchers to assist data custodians to provide researchers with de-identified research datasets in which records for individuals can be linked to records for the same individual in other datasets provided by other data custodians.

One of the early research priorities identified by stakeholders in the NT was to know more about the important factors that influence the early development of Territory children. An initial project was undertaken to link records from four sources – perinatal data register, NT Government school enrolment and attendance dataset, the Australian Early Development Census and the National Assessment Program – Literacy and Numeracy (NAPLAN). This project demonstrated the value and feasibility of linking datasets and led to an ongoing

Child and Youth Development Research Partnership
Research Brief, CYDRP Overview (June 2022)

partnership between the NT Government and Centre for Child Development and Education (CCDE) at Menzies School of Health Research. The initial focus on early childhood development was subsequently extended to include youth and, in 2018, the Child and Youth Development Research Partnership (CYDRP) was established. CYDRP is a collaboration between CCDE and seven NT Government departments.¹ From the initial linkage of four datasets to support a single project the program has gradually expanded to now contain 23 datasets in a repository that support many projects.

CYDRP data repository (Extract 3)

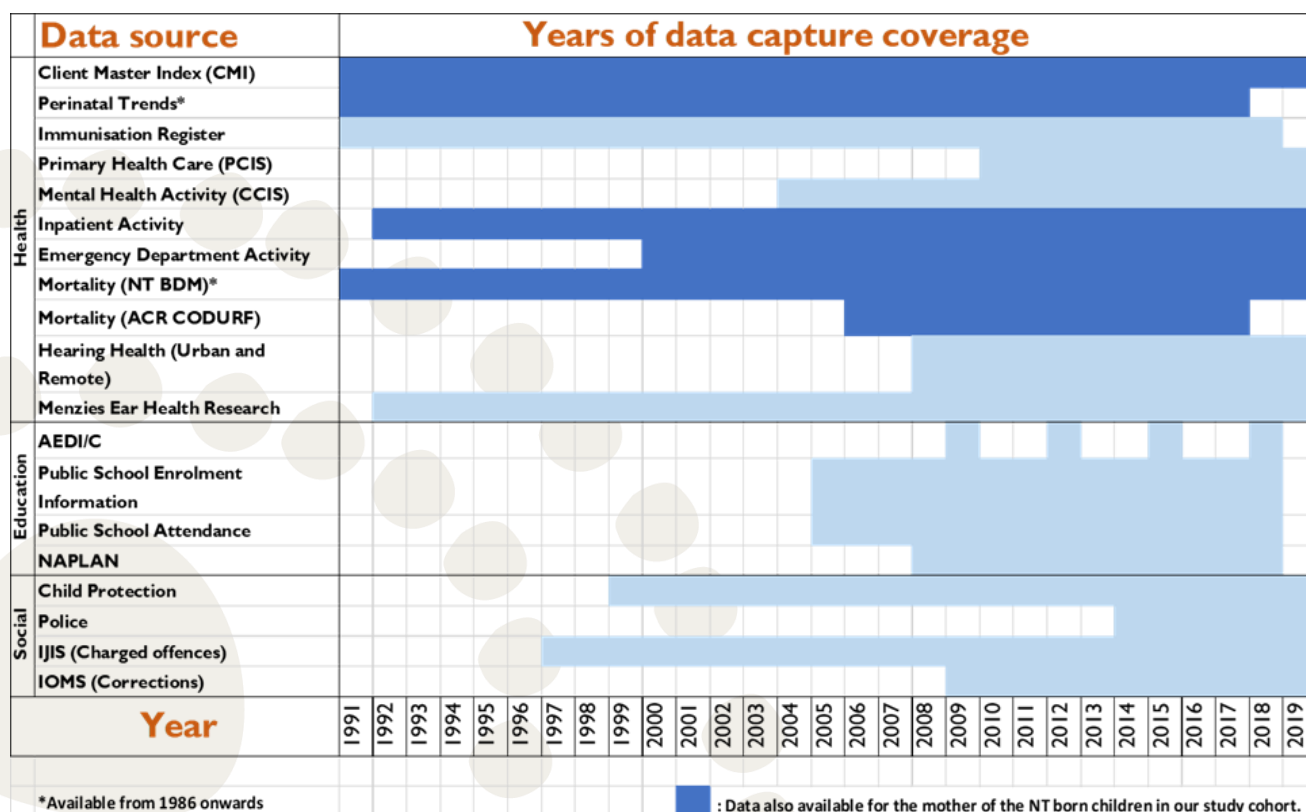
At the heart of the CYDRP program is the CYDRP data repository. The repository contains de-identified but linked information for all NT children and young people and is held on a secure server with restricted access. The repository has recently been updated, referred to as Extract 3, with records from 23 datasets across the health, education, child protection, police and justice sectors.

A summary of the datasets and timelines in Extract 3 of the repository is presented in Figure 1. The extension of the repository back to 1986 means Extract 3 contains records for young adults including young women some of whom will also have records of their own children represented in the repository. In addition, Extract 3 contains selected records for other mothers born before 1986 (highlighted in dark blue in Figure 1). There are now records for 370,000 individuals held within the repository. The update of the repository has extended the research capability to allow analysis of life-course development from birth to young adulthood, as well as assessment of the influence of characteristics of mothers and siblings on outcomes for children and young people.

What are the current projects?

There have been more than 50 projects that have utilised the repository, across a wide range of subjects including early child development, school attendance and academic outcomes, hearing impairment, child abuse and neglect, youth offending, mental health and injury.

Figure 1: CYDRP data repository (Extract 3)



¹ The NT Government CYDRP partners are: Department of the Chief Minister and Cabinet, Department of Treasury and Finance, Department of Territory Families, Housing and Communities,

Department of Education, NT Health, Department of the Attorney-General and Justice, NT Police, Fire and Emergency Services