

# LifeCourse Data Framework

Version 1.0 | Last updated July 2022



creating  
possible

# Contents

SUMMARY..... 2

BACKGROUND ..... 3

PURPOSE OF THIS FRAMEWORK ..... 4

KEY PRINCIPLES ..... 4

DAY-TO-DAY MANAGEMENT OF COHORTS ..... 5

STUDY METADATA..... 5

    Security and protection of study metadata ..... 7

    Available resources ..... 7

FACILITATING DATA ACCESS ..... 7

    Security and protection of research data ..... 8

    Available resources ..... 8

CONCLUSIONS..... 9

# Summary

The aim of the LifeCourse initiative is to enable researchers to capitalise on the availability of extensive cohort data to advance understanding of health issues emerging over the life span. To this end, LifeCourse brings together information regarding over 30 cohorts into a common platform. Each cohort is independently managed by the relevant academic or research organisation leading that study, with their own ethics and governance protocols in place. Murdoch Children's Research Institute (MCRI) operates LifeCourse and collaborates with these academic and research organisations and their respective cohort teams to promote reuse of these valuable data, including by:

- presenting cohort metadata in a standardised format on a publicly accessible website, and
- providing a central gateway for data access requests, which are evaluated and actioned by each relevant cohort data custodian.

The purpose of this document is to provide a framework for considering current data governance arrangements at the intersection between cohorts and LifeCourse, particularly regarding study metadata and data access. The primary audience are cohort data custodians and teams who undertake the day-to-day operations of their study. Articulating this framework is timely in the context of efforts across MCRI to support best practice approaches to data custodianship. We anticipate this framework will continue to evolve over time in response to changing cohort needs, Institute conventions, available technology, and best practice standards.

## Background

The Melbourne Children's LifeCourse initiative (<https://lifecourse.melbournechildrens.com>) was established in 2013, bringing together the population and clinical cohort studies (including registries) that were hosted by or in collaboration with the Murdoch Children's Research Institute (MCRI) at that time. LifeCourse has involved collaboration with a wide range of organisations such as the University of Melbourne Department of Paediatrics and the Royal Children's Hospital (which, with MCRI, comprise the Melbourne Children's Campus), as well as others via the participating cohorts and research teams. The initiative now supports over 30 cohorts, each of which is independently managed by the relevant research organisation or academic institution leading that study, with their own ethics protocols, governance models, and practices in relation to data custodianship and stewardship.

LifeCourse aims to enable local and international researchers to capitalise on the availability of these extensive cohort data to advance understanding of health issues emerging over the life span. LifeCourse undertakes a range of activities and projects to facilitate this (Figure 1). This includes standardising and organising study metadata for presentation on a publicly accessible website (<https://lifecourse.melbournechildrens.com>), allowing researchers to identify relevant data more easily within and across cohorts. Another key activity is the provision of a central gateway for data access requests, promoting a standardised, efficient and transparent application process. LifeCourse also acts as a liaison point and knowledge hub. For example, the LifeCourse team is available to consult with cohort teams about relevant supports and resources available across the Campus that support best practice in cohort management.

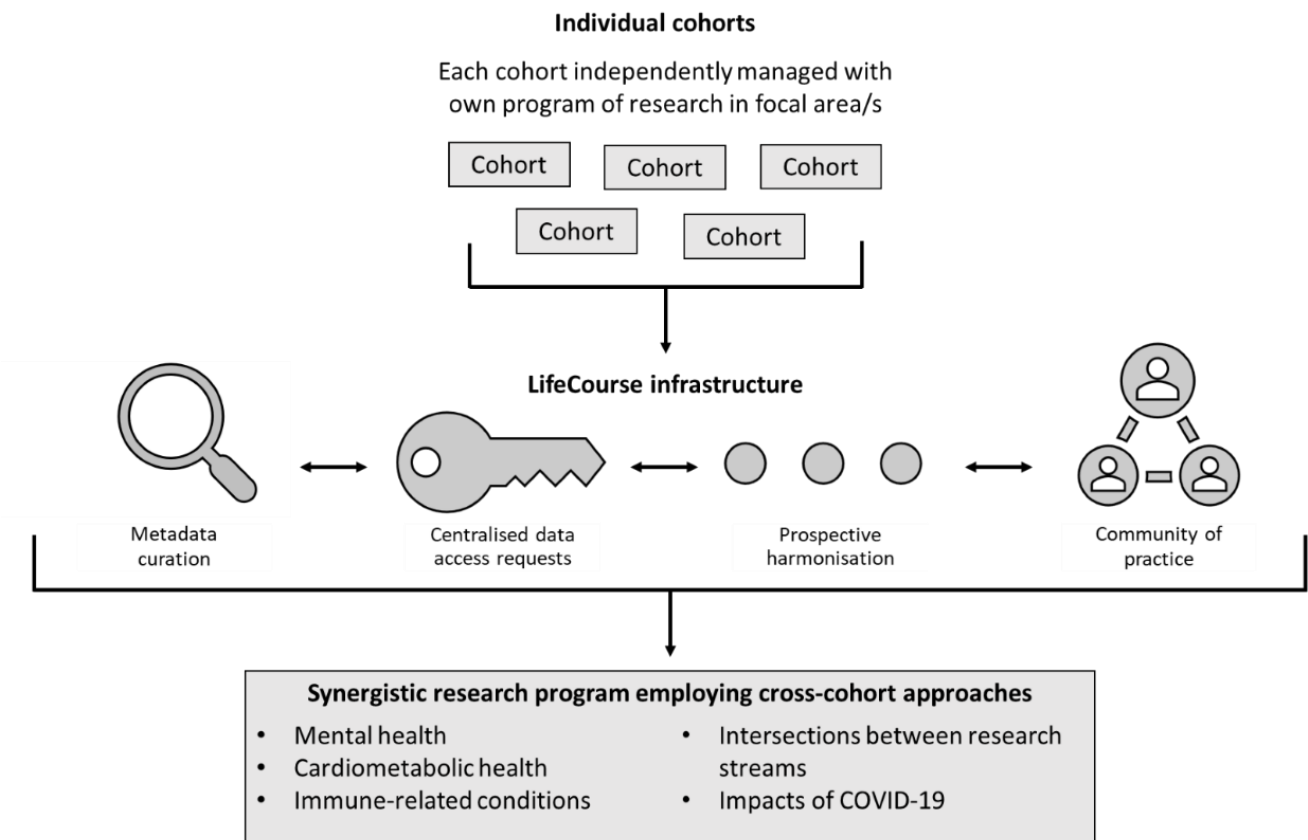


Figure 1. Overview of the LifeCourse initiative.

## Purpose of this framework

Cohort data and metadata are independently managed by each cohort's study team with their own ethics and governance protocols, and with cohort investigators acting as the data custodians. LifeCourse similarly has its own governance protocols and reporting structures. The purpose of this document is to clarify current data governance arrangements at the intersection between the cohorts and LifeCourse, including use of study metadata and roles and responsibilities in relation to data access. Articulating this framework is timely in the context of efforts by the Institute to support best practice data management and governance, which is currently an area of focus and activity. Other key Campus groups that support cohorts in issues relating to data governance include the Royal Children's Hospital (RCH) [Research Ethics and Governance \(REG\)](#) office, the Office of Research at MCRI, and the Centre for Epidemiology and Biostatistics Unit (CEBU) at MCRI.

## Key principles

The LifeCourse infrastructure is informed by FAIR (Findable, Accessible, Interoperable, Reusable) data principles that maximise the value of existing data assets. Complimentary to this is a commitment to Open



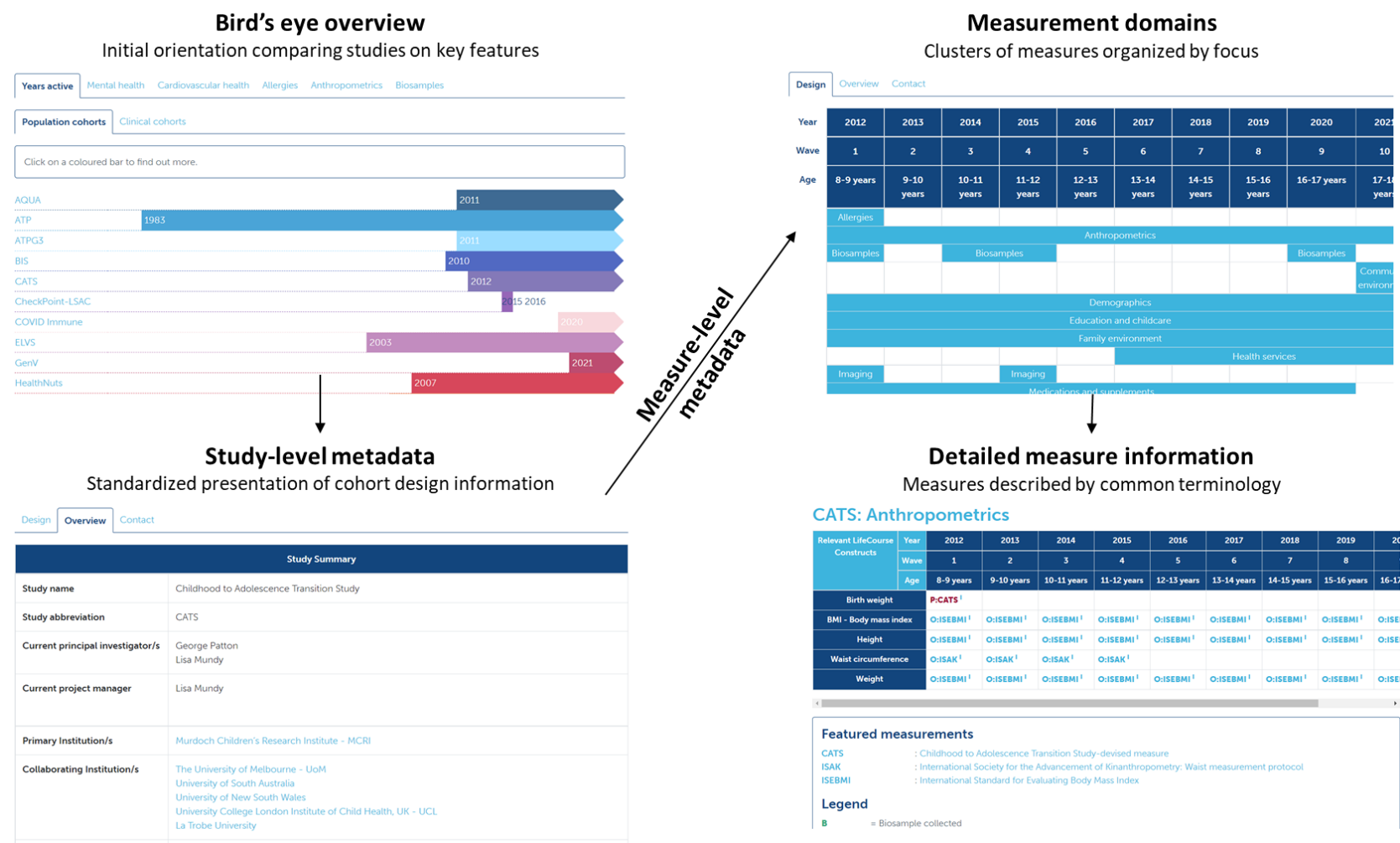
Science, and the promotion of science that is transparent and reproducible. Facilitating team science is also fundamental. LifeCourse is committed to building a community of practice across cohort researchers, who can deepen their collective expertise by sharing knowledge, approaches, and insights on common problems related to cohort management and data stewardship.

## Day-to-day management of cohorts

Cohort custodians are responsible for the day-to-day management and custodianship of their cohort, in line with their governance arrangements and ethical approvals. This includes responsibility for activities such as strategic planning, ethical oversight, submission of funding proposals, development of study documentation, and management of data sharing. LifeCourse collaborates with custodians to facilitate increased consistency and use of best practice approaches across such areas. It does not overtake or conflict with cohorts' existing governance arrangements. Rather, it supports cohort custodians through a community of practice approach to increasing communication, consistency, and cooperation.

## Study metadata

To maximise discoverability of these participant data for researchers, the LifeCourse platform collates standardised study metadata for presentation on a publicly accessible website. Study metadata are organised at several levels (Figure 2). At the highest level, description of key design features, such as the year established and number of participants, is provided in a standardised format. At a more detailed level, measures captured within each wave of data collection are described according to a common terminology and organised into content domains. Using this standardised system of description provides consistency with international standards and facilitates external comparisons of data availability.



Responsibility for the preparation and collation of study metadata to feed into the LifeCourse website sits with individual cohorts, who have access to their full study documentation and know their study best. Cohorts are responsible for the accuracy of the collated metadata provided to MCRI through LifeCourse, and for notifying LifeCourse of updates required to this metadata as new waves of data collection are progressed. LifeCourse is responsible for organising these metadata into a browsable and searchable format for presentation on the publicly accessible LifeCourse website hosted on the MCRI server.

Metadata presented on the LifeCourse website is in no way intended as a substitute for each study's own detailed documentation and codebooks, which remain the responsibility of individual cohorts. The LifeCourse metadata catalogue provides an entry point for data users to easily understand what data are available within and across cohorts, which would otherwise be highly onerous, resource demanding, and often not practically feasible for external researchers. As a point of first access, the metadata presented on the LifeCourse website is not comprehensive and does not include all information required during later analysis planning, such as the specific items contained within scales, question wording, or scoring approaches.

## Security and protection of study metadata

LifeCourse stores cohort's raw metadata in MCRI's secure cloud storage. This information is formatted and organised by LifeCourse for presentation on the publicly accessible LifeCourse website. Any associated cohort study documentation (e.g., data dictionary, survey forms) provided to LifeCourse are also stored in MCRI IT's approved storage systems, and will not be shared by LifeCourse with any third parties without cohort consent.

## Available resources

Resources available to assist cohorts in the development of their study documentation are detailed at: <https://lifecourse.melbournechildrens.com/#data-management-and-analysis>. In particular, the Clinical Epidemiology and Biostatistics Unit (CEBU) is available to provide information, advice and guidance on managing data (policies, planning, handling data), data management software, database setup, and data capture. The LifeCourse team are available to consult on data documentation practices that allow for easy extraction of required information for the LifeCourse metadata catalogue.

## Facilitating Data access

To reduce logistical barriers to data access, LifeCourse acts as a liaison connecting data users and custodians. LifeCourse manages the application process administratively, acting as an intermediary between the applicant and custodian. Cohorts are responsible for deciding whether to progress the



application at each stage and for actioning approved projects. An application proceeding through the LifeCourse system should not be interpreted as providing LifeCourse's endorsement of the proposed project, as this automated system does not include a review or evaluation by the LifeCourse team.

To initiate the process, applicants are invited to complete an initial enquiry

(<https://lifecourse.melbournechildrens.com/data-access/>), requiring preliminary information on the team, primary research question, and cohort/s of interest. LifeCourse confirms with the relevant cohort/s whether they would like to proceed with the request. If so, an application covering full details of the project and data and/or samples required is submitted by the applicant. If cohorts provide in-principle approval for the project, LifeCourse introduces the applicant and custodian to begin their direct communication around next steps. From there, cohorts are responsible for ensuring that any resulting data access aligns to their ethical consents and approved processes, including entering into appropriate written agreements with applicants. MCRI through LifeCourse tracks the research outputs arising from the request (e.g., papers published), to monitor and improve the system as well as provide key metrics back to cohorts.

Cohort custodians retain decision-making responsibility and undertake the transfer of data and/or samples on their own terms and conditions. Applications are assessed by each relevant research or academic organisation responsible for the cohort on criteria such as:

- feasibility given the available data (e.g., quality issues with the data requested)
- consistency with ethical requirements (e.g., limits of participant consents)
- appropriateness for the purpose and strategic plans of the cohort (e.g., redundancy with research already underway), and
- scientific quality.

MCRI is not involved in the contracting process except when MCRI is the relevant party providing access to the data. The relevant research organisation or academic institution responsible for the cohort enters into a separate agreement with the applicant to cover the transfer and use of data.

## Security and protection of research data

LifeCourse does not hold or have direct access to any individual participant data. Cohorts are responsible for the management of their data and ensuring that any data sharing aligns to their ethical consents and approved processes. Ownership of all participant data and raw metadata remains with the cohort and their relevant organisations.

## Available resources

Cohorts can find further information and resources that support appropriate data sharing at:

<https://lifecourse.melbournechildrens.com/#data-management-and-analysis>. The [RCH REG](#) office are

available for consultation on ethical issues relating to data reuse. The LifeCourse team can consult on how to integrate use of the LifeCourse data access request portal into a broader data sharing plan.

## Conclusions

The Melbourne Children's LifeCourse Initiative enables researchers to more effectively leverage the value of existing cohorts to improve child and adolescent health. Beyond the quality, scope and richness of the underlying cohorts, strengths of the LifeCourse initiative include the availability of richly described and structured cohort metadata and a common approach to data access requests. This provides efficiency, comparability, and feasibility in the use of these data, enhancing their value for promoting life course health. This framework provides an overview of how LifeCourse and individual cohorts currently collaborate to further these goals while maintaining the highest standards of data security and ethical practice. We anticipate this framework to evolve as LifeCourse continues to strengthen and change over time in response to cohort needs, Institute conventions, available technology, and best practice standards. Such further developments will be captured in updated releases of this framework.