Fás Aníos in Éirinn
Growing Up in Ireland

# Scoping Review to Inform the Development of a Potential New Birth Cohort for Growing Up in Ireland

May 2023

An Roinn Leanaí, Comhionannais, Míchumais, Lánpháirtíochta agus Óige
Department of Children, Equality, Disability, Integration and Youth

University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Department of Children, Equality, Disability, Integration and Youth
Block 1, Miesian Plaza, 50 – 58 Lower Baggot Street, Dublin 2
D02 XW14
Tel: +353 (0)1 647 3000
Email: research@equality.gov.ie
Web: www.gov.ie/dcediy

# Scoping Review to Inform the Development of a Potential New Birth Cohort for Growing Up in Ireland

Dearbháile Counihan*, Marie T. O'Shea*, Daniel Guigui, Seaneen Sloan,
Yekaterina Chzhen+, Laura K. Taylor+, Jennifer E. Symonds+

*Joint first authors +Joint senior authors

University College Dublin, Geary Institute for Public Policy
University College Dublin, School of Psychology
University College Dublin, School of Education
Trinity College Dublin, School of Social Sciences and Philosophy, Department of Sociology

This report was funded by the Department of Children, Equality, Disability, Integration and Youth. The authors are solely responsible for the views, opinions, findings, conclusions and/or recommendations expressed, which are not attributable to the department.

# Contents

# List of Tables

# Acknowledgements

The research team have several acknowledgements to make. First, they would like to thank the advisory group for their support and assistance with the scoping review. The advisory group were Prof. Richard Layte (Trinity College Dublin), Prof. Dermot O'Reilly (Queens University Belfast), Dr. Orla McBride (Ulster University), and Prof. Lisa Calderwood (University College London). Second, the research team express their sincere thanks to the participants in the stakeholder consultation who freely gave their time to discuss the context surrounding a potential new Growing Up in Ireland birth cohort. Third, the research team thank the study teams who communicated with them about their cohort studies featured in this scoping review. Finally, the research team would like to acknowledge the intellectual, creative, and editorial contributions of the Department for Children, Equality, Disability, Integration, and Youth; especially Anna Visser and Claire Farrell. The quality of this Scoping Review is testament to the involvement of all people who contributed across sectors

# List of acronyms

| Acronym | Definition |
|---|---|
| AIHW | Australian Institute of Health and Welfare |
| ALSPAC | Avon Longitudinal Study of Parents and Children |
| AMF | Anonymised Microdata File |
| BAS | British Ability Scales |
| BMI | body mass index |
| CAO | Central Applications Office |
| CAPI | computer-assisted personal interviewing |
| CASI | computer-assisted self-interviewing |
| CATI | computer-assisted telephone interview |
| CESSDA | Consortium of European Social Science Data Archives |
| CLS | Centre for Longitudinal Studies |
| CLSI | Centre for Longitudinal Studies in Ireland |
| CNAF | Caisse Nationale d'Allocations Familiales (French National Family Allowance Fund) |
| CNIL | Commission Nationale de l'Informatique et des Libertés (National Data Protection Authority) |
| COORDINATE | COhort cOmmunity Research and Development Infrastructure Network for Access Throughout Europe |
| COVID-19 | coronavirus disease 2019 |
| CSO | Central Statistics Office |
| DCEDIY | Department of Children, Equality, Disability, Integration and Youth |
| DEIS | Delivering Equality of Opportunity in Schools |
| DKK | Danish krone |
| DNA | deoxyribonucleic acid |
| DNBC | Danish National Birth Cohort |
| DPC | Data Protection Commission |
| DWP | Department for Work and Pensions |
| EBV | Epstein-Barr virus |

| Acronym | Definition |
|---|---|
| EHR | electronic health record |
| ELC-FS | Early Life Cohort- Feasibility Study |
| ELF | Early Life Factors study |
| ELFE | Étude Longitudinale Française depuis l'Enfance |
| ESRC | Economic and Social Research Council |
| ESRI | Economic and Social Research Institute |
| EU | European Union |
| EUR | euro |
| FAIR | findable, accessible, interoperable, and reusable |
| GDPR | General Data Protection Regulation |
| GenV | Generation Victoria |
| GP | general practitioner |
| GUI | Growing Up in Ireland |
| GUIDE | Growing Up In Digital Europe |
| HEI | higher education institution |
| HIQA | Health Information and Quality Authority |
| HIS | health information system |
| HSE | Health Service Executive |
| IgG | immunoglobulin G |
| IHI | individual health identifier |
| INED | L'Institut national d'études démographiques (French Institute for Demographic Studies) |
| INSEE | Institut national de la statistique et des études économiques (National Institute of Statistics and Economic Studies) |
| INSERM | Institut national de la santé et de la recherche médicale (National Institute of Health and Medical Research) |
| InVS | Institut de veille sanitaire (French Institute for Public Health Surveillance) |
| ISS | Irish Statistical System |
| ISSDA | Irish Social Science Data Archive |

| Acronym | Definition |
|---|---|
| IT | information technology |
| LSAC | Longitudinal Study of Australian Children |
| MCS | Millennium Cohort Study |
| mHealth | mobile health |
| MN-CMS | Maternal and Newborn Clinical Management System |
| MoBa | Norwegian Mother and Child Cohort |
| NCS | National Children’s Study |
| NGO | non-governmental organisation |
| NHS | National Health Service |
| NIH | National Institutes of Health |
| NINFEA | Nascita e INFanzia: gli Effetti dell’Ambiente |
| NIPT | non-invasive prenatal testing |
| NPRS | National Perinatal Reporting System |
| ONS | Office for National Statistics |
| PCB | polychlorinated biphenyl |
| PIN | personal identification number |
| PPSN | Personal Public Service Number |
| RMF | Researcher Microdata File |
| RR | relative risk |
| SAIL | Secure Anonymised Information Linkage |
| SAPS | Small Area Population Statistics |
| SNIIRAM | Système National d’Informations Inter-Régimes de l’Assurance Maladie (France’s social security [data management] system) |
| SPSS | Statistical Package for the Social Sciences |
| UK | United Kingdom |
| USA | United States of America |
| VDI | virtual desktop infrastructure |
| WoS | Web of Science |

# Chapter 1: Introduction

## 1.1 Introduction

The health and well-being of adults is established in early life. A nationally representative birth cohort study that begins during pregnancy or from birth has a unique value in understanding the health and well-being processes of children into adulthood (Canova & Cantarutti, 2020; Golding, 2009). The Growing Up in Ireland (GUI) national longitudinal study has to date played an important role in understanding the varied lives of children in Ireland.

This scoping review will assess the scope of a potential new GUI birth cohort by examining developments across key domains of data collection and analysis for longitudinal birth cohort studies, in addition to exploring relevant health, socioeconomic and environmental factors across the life course of the next cohort of children in Ireland. Significant social, economic and policy changes have occurred in Ireland (e.g. population composition, healthcare, education, labour market) since the beginning of the GUI birth cohort study in 2008/2009. Recent policy changes also include those brought about by the COVID-19 pandemic and the war in Ukraine.

## 1.2 Chapter overview

This chapter will begin with an overview of the background to longitudinal birth cohort studies and how they can help establish causal links between exposures and outcomes. This overview will then lead into a literature review to synthesise the conceptual frameworks used in birth cohort studies, providing a synopsis of a range of study objectives across multiple international birth cohort studies. The selection process used to determine the four named birth cohort studies to serve as case studies is outlined. These cases are the Millennium Cohort Study (MCS), the Danish National Birth Cohort (DNBC), Étude Longitudinale Française depuis l'Enfance (ELFE), and Generation Victoria (GenV). Each case study is then addressed with respect to the study background, rationale, aims and objectives, conceptual framework, main study domains, and sampling framework. However, more detailed discussions of the data collection methods, study designs (including attrition), and data access and analysis are provided in Chapters 2, 3 and 4, respectively. The Early Life Cohort Feasibility Study (ELC-FS)[1] and the Growing Up In Digital Europe (GUIDE)[2] study are then discussed, although not as named case studies for the purpose of this review. The ELC-FS is a new United Kingdom (UK)-based birth cohort study that will offer valuable insights into a range of learnings to benefit future birth cohort study research teams. GUIDE is Europe's first cross-country comparative birth cohort study. Finally, a discussion of the strengths and limitations of study conceptual frameworks concludes this chapter, keeping in mind the GUI '08 study as the 'foundation case'.

## 1.3 Background to longitudinal birth cohort studies

Infancy and childhood are critical life phases typified by rapid growth and development processes which influence health and well-being across the life span. Early-life exposures, including nutritional, environmental, and socioeconomic factors, can affect optimal growth and development from prenatal stages and throughout the life span (Larsen *et al.*, 2013; Lawlor, Andersen & Batty, 2009; Lynch & Smith, 2005). Broader political, societal, and cultural contexts also affect infancy and early childhood development phases (Black *et al.*, 2017; Maggi *et al.*, 2010). Longitudinal birth cohort studies have contributed substantially to increasing scientific knowledge about prenatal, childhood, and life course health outcomes (Orri *et al.*, 2020; Ji *et al.*, 2019; Thompson *et al.*, 2010; Golding, 2009; Burke *et al.*, 2005). Birth cohorts can also shed light on socioeconomic, racial and ethnic, and other structural disparities across a range of social, health, and economic outcomes (Cotter *et al.*, 2019; Kent & Pitsia, 2018; McCrory *et al.*, 2017; Watson *et al.*, 2014; Kelly, Becares & Nazroo, 2013; Sullivan, Ketende & Joshi, 2013).

Longitudinal birth cohort studies are prospective studies which follow the same cohort of participants over time from birth. Multiple domains of data are collected to provide insights into the complex processes and outcomes across participants' lives (Canova & Cantarutti, 2020; de Groot *et al.*, 2017). Prospective designs are key to explaining and understanding the causal direction of an association by enabling the

---

[1] For additional information, visit the ELC-FS study website: https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study/

[2] For additional information, visit the GUIDE study website: https://www.guidecohort.eu/

identification of predictors and a range of factors (both risk and protective) that are associated with outcomes and/or trajectories/pathways over time (Richmond *et al.*, 2014). From a social health perspective, longitudinal data can help uncover the cause and effect of inherited and life course influences on childhood health, and the development of common and complex health issues into adulthood. While many birth cohort studies are epidemiologically oriented, many recent birth cohort studies (such as the MCS, Growing Up in Scotland, Growing Up in New Zealand, and Growing Up in Australia: The Longitudinal Study of Australian Children [LSAC]) study the interplay of factors across multiple domains. Furthermore, the long-term follow-up period in birth cohort studies makes them valuable tools to address new policy-related questions and influence policy development (Checkoway, Pearce & Kriebel, 2007). Birth cohort studies can also assess the impact of policy changes over time. For example, children in the GUI study were among the first to be eligible for a subsidised preschool place arranged through the Early Childhood Care and Education scheme (Smyth, 2018). This allowed for the assessment of this new policy regarding take-up and the multiple effects of centre-based care on children.

Since birth cohort studies follow the same individuals over time, they can help establish causal links between exposures and outcomes. Exposures and events in the pre- and postnatal periods and infancy can result in long-term consequences for health and well-being into adulthood. For example, prenatal development and birth outcomes can be influenced by maternal diet, drug use, smoking, and alcohol intake (Heude *et al.*, 2016: Lawlor, Andersen & Batty, 2009). Early life factors and exposures have been associated with cardiovascular disease, cancer, asthma, other respiratory diseases, allergies, cognitive ability and degeneration, socio-emotional function, and mental health (Mitku *et al.*, 2021; Blane *et al.*, 2013; Hanson & Gluckman, 2011; Gluckman *et al.*, 2008). Exposure to poverty and/or material deprivation are additional risk factors for poor health in childhood (Marmot & Wilkinson, 2005).

# 1.4 Literature review of core conceptual frameworks and study objectives

## 1.4.1 Overview of conceptual frameworks for longitudinal birth cohort studies

Conceptual frameworks underpin research by providing a set of definitions or approaches to follow (Corna, 2013). Used frequently in social, health and behavioural science research, conceptual frameworks include one or more theories and concepts to illustrate causal pathways and relationships, and how they map onto the observational or interventional studies being undertaken. In social and health-focused research, conceptual frameworks guide intervention studies and offer insights into how to approach reducing disparities (Ridgeway *et al.*, 2017).

Longitudinal birth cohort studies often employ conceptual frameworks based on life course models (Jones *et al.*, 2019; Kuh *et al.*, 2003). Life course conceptual frameworks are interdisciplinary and are used in the domains of sociology, biology, psychology, anthropology, economics, and epidemiology (Jones *et al.*, 2019; Kuh *et al.*, 2003). These frameworks guide understanding of the short- and long-term effects of a myriad of social, health and economic factors which are affected by a range of exposures and determinants from conception to death (Jones *et al.*, 2019). These range from biomedical, socioeconomic, and psychosocial factors to environmental factors at macro, meso, and micro levels, and take period effects taken into consideration (e.g. children born at the end of the 1990s in comparison with children born in the 2000s) (Greene *et al.*, 2010a; Kelly *et al.*, 2009).

Wang *et al.* (2021) reviewed a comprehensive range of conceptual frameworks used by early life cohort studies worldwide. The purpose of this review was to inform the selection and development of a relevant life course framework to guide the content selection and visual communication design for the GenV study. GenV is a new longitudinal research programme based in Victoria, Australia. This health-focused, large-scale cohort study aims to examine and address complex health issues affecting children from birth to adulthood (Davies *et al.*, 2020). GenV will follow a cohort of children born in Victoria between mid-2021 and mid-2023. The comprehensive range of frameworks identified by Wang *et al.* (2021) are summarised in Table 1.1. The conceptual frameworks used in the selected case studies for this report are discussed later in this chapter (see Section 1.4 for further detail).

**Table 1.1** Conceptual frameworks related to early life cohorts.

| | |
|---|---|
| **1.** | Biodevelopmental framework (Shonkoff, 2010): a combined biodevelopmental framework to enhance improved interpretation of the antecedents and causal pathways that lead to inequalities in health, development, learning, and behaviour. It aims to advance theories of change in order to enhance policies and programmes. |
| **2.** | Life Course Health Development (Halfon *et al.*, 2014): a transdisciplinary framework from a developmental psychology and biology perspective to harness multiple theories to increase understanding of how health develops and changes over the life course. |
| **3.** | Biosocial approach to human development, behaviour, and health across the life course (McDade & Harris, 2018): a 'biosocial approach' where biological and social concepts, models, and methods are intertwined to understand the mechanisms and pathways through which socioeconomic and psychosocial factors shape health and development over the life course. |
| **4.** | Life course approach to chronic disease epidemiology (Ben-Shlomo & Kuh, 2002): a framework to understand chronic disease epidemiology which illustrates how biological, behavioural, and psychosocial exposures during the prenatal period through to adulthood can have long-term influence on the risk of developing chronic disease. |
| **5.** | European Union (EU) LifeCycle (Jaddoe *et al.*, 2020): developed by the EU Child Cohort Network to improve life course health pathways by focusing on early-life exposures. This conceptual model brings together 19 pregnancy and childhood cohorts of more than 250,000 children and their parents. |
| **6.** | Australian Institute of Health and Welfare (AIHW) person-centred framework (Australian Institute of Health and Welfare, 2021): an Australian population-focused framework to guide child well-being. Based on a social-ecological model of health, this framework studies the domains of health and welfare for the general population. |
| **7.** | Murdoch Children's Research Institute (MCRI) complex disorders framework: a complex disorders model which recognises the effects of environmental exposures and individual characteristics on the 'bio-signature' of each individual and uses data to predict and prevent complex disease. |
| **8.** | Life course model of ageing (Hanson *et al.*, 2016): a framework that views ageing as an adaptive procedure where the risks of health problems accumulate from embryo development throughout the life course and can be passed from one generation to the next. Risks for many public health problems, including non-communicable diseases, accumulate throughout an individual's life from the embryo stage onwards. |

| | |
|---|---|
| **9.** | The National Institutes of Health's (NIH) Environmental Influences on Child Health Outcomes (ECHO) model of positive health (Forrest, Blackwell & Camargo, 2018): a model which theorises that individual health is shaped by multifocal environmental exposures which occur repeatedly over the life course. These include the social, family, and physical interactions that influence health. |
| **10.** | Pathways to intrauterine growth retardation (Spencer, 2010): a framework to explore how social inequalities in child health can impact on intrauterine growth retardation over time and across generations. |
| **11.** | New Zealand's living standards framework (New Zealand Treasury, 2020): a New Zealand-based framework to increase understanding of how to improve living standards to enhance well-being across generations. |
| **12.** | Bioecological model (Bronfenbrenner, 1979): a widely used framework that explores child development within the context of the individual's interactions across multiple environments (e.g. family, school, cultures) and at different levels of the social environment, including the microsystem, mesosystem, exosystem, macrosystem, and chronosystem. |
| **13.** | Resilience models in life course (Cosco *et al.*, 2017): an approach that integrates methods to determine how challenging events and positive changes can be harnessed to operationalise the concept of resilience. |
| **14.** | Multilevel factors related to paediatric health disparities (Ridgeway *et al.*, 2017): a framework developed via a systematic review to understand the multiple factors related to paediatric health disparities. |

Source: Wang *et al.*, 2021, pp. 3–5

Other primarily epidemiologically focused conceptual frameworks are the Fetal Origins Hypothesis, also known as the Barker Early Origins Hypothesis (Barker, 1995); the Developmental Origins of Health and Disease (DOHaD) hypothesis, a hybrid developed from the Fetal Origins Hypothesis (Gluckman, Hanson & Buklijas, 2010; Barker, 2007); chronic disease prevention frameworks (Raine *et al.*, 2006; WHO, 2006); and a social reciprocal framework (Knight *et al.*, 2009).

### 1.4.2 Overview of study objectives for longitudinal birth cohort studies

There are multiple ongoing or recently concluded birth cohort studies with a diverse range of study objectives. In the European context, there are numerous prospective and retrospective birth cohort studies commencing at different stages, collecting data at different time points, and incorporating different rationales, aims and objectives specific to the study's country (Pansieri *et al.*, 2020). Several of the studies identified below are not recent birth cohort studies, and are not nationally representative (e.g. Born in Bradford). However, they are valuable examples of the range and breadth of birth cohort studies' focus and domains.

Some cohort studies are broad and multidisciplinary (e.g. GUI, Growing Up in Scotland, Growing Up in New Zealand, Growing Up in Australia: The LSAC, the MCS). These studies have a broad range of aims and objectives which are not singularly health focused. For example, Growing Up in New Zealand was established to capture data (with the potential to influence policy development) from a multidisciplinary perspective about the health and development of children born in New Zealand (Morton *et al.*, 2013). The study focused on identifying how influencing factors (including political, social, cultural, intergenerational, and family relationships) affect children across the life course. Growing Up in Australia: The LSAC also examines the influence of family, socioeconomic and cultural environments on child development and well-being, with the primary aim of influencing and improving child-focused policies and developing intervention and prevention strategies to improve the lives of children and their families (Gray & Sanson, 2005). The study focuses on the mechanisms that influence child well-being and development and how these mechanisms change over the life course. Similarly, Growing Up in Scotland's principal aim is to collect evidence on the lives of children through to adulthood in order to support policy creation within a Scottish context (Anderson *et al.*, 2007). The study focuses on social influences, behavioural development, physical and mental health, family life, education, and employment (Bromley & Cunningham-Burley, 2010). Likewise, the GUI study aims to inform government policy pertaining to children, young people, and their families through an exploration of children's health, education, and cognitive and socio-emotional development within varied social, economic, and cultural contexts (Growing Up in Ireland, 2022).

A large body of birth cohort studies are more epidemiologically focused. Many of these studies have a wide range of aims and objectives related to collecting data on maternal health, environmental factors, and other factors that impact on the perinatal and early life health of children (e.g. BAMSE in Sweden (Wickman *et al.*, 2002), the EDEN study in France (Heude *et al.*, 2016), and the Polish Mother and Child

Cohort Study (REPRO_PL study) in Poland (Polańska *et al.*, 2011)). The EDEN study is a longitudinal epidemiological birth cohort study with the primary objective of gathering evidence on the pre- and early postnatal determinants of child health and development (Heude *et al.*, 2016). The study is heavily focused on environmental factors and exposures that may affect child health, such as maternal nutrition and its influence on obesity and asthma. It also collects data on socioeconomic factors and their effects on the biological health of the child. Generation R (Netherlands) is a multidisciplinary study whose primary objective is to better understand early-stage environmental and genetic causations of irregular growth, development, and child health from prenatal stages through to adulthood (Kooijman *et al.*, 2016). The study is health focused, with an emphasis on a wide range of health outcomes from endocrine, genomic, nutritional, environmental, and socioeconomic determining factors.

The primary objectives of the Avon Longitudinal Study of Parents and Children (ALSPAC) in the UK were to investigate the multiple ways in which genetic and environmental attributes influence health and development of children from early life into adulthood (Golding *et al.*, 2001). This study was particularly focused on a specific area in Southern England. Another UK-based study, Born in Bradford, had broad objectives to compare health and well-being within a bi-ethnic population and to investigate the causal pathways that either promote health or contribute to poor health (Raynor & Born in Braford Collaborative Group, 2008). This work was undertaken with the aim of developing a model for linking and incorporating research into routine data administrative systems in the National Health Service.

Frequently addressed in birth cohort studies are the exposures which may affect the development of child health outcomes, including asthma, allergy, weight and height, mental health, neurodevelopment, and infectious disease. For example, the LIFE Child (Germany) study’s primary objective is to better identify the growth and development of diseases such as obesity through the monitoring of children from birth into adulthood (Poulain *et al.*, 2017). The study utilises a three-cohort design to understand the differences between a birth cohort, a health cohort, and an obesity cohort. The REPRO_PL study aims to understand the impact of exposures to differing environmental factors during prenatal and postnatal stages, and to measure their effect on pregnancy and child health outcomes (Polańska *et al.*, 2011). The study particularly focuses on toxins and their potential impact on the development of respiratory diseases, allergy, and delayed mental and physical well-being.

Some birth cohort studies also incorporate novel technologies to study child health. For example, the KUNO-Kids birth cohort study was established with three primary objectives. First, to contribute to the body of research into child health using novel technologies, with an interdisciplinary systems medicine approach, and to use data to the benefit of the individual. Second, to identify possible prevention strategies for improved child health. And third, to create an outline to approach examining the feasibility and effectiveness of targeted health interventions (Brandstetter *et al.*, 2019).

Growing numbers of studies collect biological samples (e.g. maternal blood, maternal serum/plasma, maternal deoxyribonucleic acid (DNA), breast milk, child blood, child serum/plasma, child DNA, umbilical cord blood, paternal DNA) and/or establish biobanks (e.g. the Aarhus Birth Cohort Biobank (Denmark), DNBC (Denmark), ELFE (France), and the Multiple Birth Cohort Study (MUBICOS) (Italy)). Often, these studies have an objective to investigate the genetic determinants of disease or biomarkers of certain exposures in order to assess their potential impact on prenatal and early life health. However, biobanks require complex funding and long-term investment which can be financially restrictive for birth cohort studies (Doyle & Golding, 2009).

A comprehensive overview of a variety of birth cohort studies' objectives and data topics, including at different waves and different ages of the cohort participants, can be found on the website Birthcohorts.net.[3] The purpose of this website is to serve as a repository of information about the design of, and data collected by, birth cohort studies internationally, and to facilitate knowledge exchange, improve study collaboration, and increase information accessibility for researchers, policy-makers and other stakeholders. It is important to note that while this website identifies many longitudinal birth cohort studies, it is not an exhaustive repository of all such studies worldwide.

---

[3] See https://www.birthcohorts.net/ for further information.

## 1.5 Learning from existing birth cohort studies

To inform this report, Cohort '08 of the Growing Up in Ireland (GUI) study is used as the foundation case study – the main point of reference for explorations of other international birth cohorts. In addition to the GUI study, this report also includes an in-depth examination of four selected longitudinal birth cohort studies that exhibit innovative techniques internationally.

### 1.5.1 Case study selection: decision-making criteria

Studies were identified using strict inclusion and exclusion criteria in order to identify the longitudinal birth cohort studies that were current and relevant. Studies were deemed eligible if they: (a) focused on infants under 1 year of age at the time of the first data collection, (b) were conducted in the last 20 years, (c) were nationally or regionally representative, and (d) utilised a large sample. Additional criteria used to narrow down the potential birth cohort studies included: (e) taking measurements at multiple time points (longitudinal versus only early childhood) and (f) using diverse forms of data (e.g. surveys and biological samples) and/or linking to population registries (e.g. biobanks, school or neighbourhood data).

### 1.5.2 Selection process

The freely accessible database Birthcohorts.net[4] was used to locate potential named birth cohort studies from around the world. Of the 134 birth cohort studies accessible from this database, alongside the additional longitudinal birth cohort studies found in consultation with the research team, a short list was compiled of 23 studies that met our eligibility criteria (Appendix 1A). The list of potential studies was reviewed concerning date (most recent), sample size (representative), follow-up (i.e. not only early life/infancy), coverage (domain and topics), and data diversity (e.g. biomedical data, linked data). In discussion with the Department of Children, Equality, Disability, Integration and Youth (DCEDIY) and our Advisory Board (comprising of researchers with significant expertise in the field), four named birth cohort studies were selected to serve as case studies alongside Cohort '08 of GUI: the MCS, DNBC,

---

[4] The complete inventory of cohorts listed on the Birthcohorts.net website did not include the GUI study or the MCS. Cohorts are registered to this website by the researchers themselves.

ELFE, and GenV studies. The key features of these studies are summarised below. See Appendix 1A for a comparison of the study features across the selected case studies. A brief overview of the newly launched ELC-FS (UK) is also provided later in this chapter (see Section 1.5.8 The Early Life Cohort Feasibility Study).

## 1.5.3 GUI Cohort ’08: the foundation case

### 1.5.3.1 Background

GUI is the national longitudinal study of children in Ireland, funded by DCEDIY, and managed by DCEDIY in association with the Central Statistics Office (CSO) (Greene *et al.*, 2010a; Greene *et al.*, 2010b). The study is conducted on behalf of DCEDIY by a team at the Economic and Social Research Institute (ESRI) in collaboration with researchers at Trinity College Dublin. GUI is the largest and most comprehensive study of its kind in Ireland. It began in 2006 and follows two cohorts of children: Cohort ’98 (formerly called the Child Cohort), who were aged 9 years at the first data collection phase in 2007, and Cohort ’08 (formerly called the Infant Cohort), who were aged approximately 9 months at the first wave of data collection in 2008. For this report, GUI Cohort ’08 is the foundation case study. Phase 1 covered data collection at ages 9 months, 3 years, and 5 years of the Infant Cohort ’08. Phase 2 (2015–2019) covered ages 7/8 years and 9 years for the Infant Cohort ’08. An extension was added to Phase 2, which covers a fifth wave of the Infant Cohort ’08 at age 13 years in 2021.

GUI participants and their parents also took part in a COVID-19 web survey in 2020. GUI Cohort ’08 participants were aged approximately 12–13 years at the time of data collection. The survey aimed to provide insights into their experience of the social, health, and economic effects of the COVID-19 pandemic (e.g. physical and mental health, family environment, home education, Internet use, emotional well-being, diet and exercise) (Kelly *et al.*, 2021).

### 1.5.3.2 Rationale

The initial call for a national longitudinal children’s study in Ireland dates back to the 1980 Task Force on Child Care. However, it was not until the 2000 National Children’s Strategy that the Irish government announced its commitment to commissioning such a study. The study was commissioned in 2005/06 following work on the design and conceptual framework. *The National Children's Strategy: Our Children – Their Lives* (Department of Health and Children, 2000) identified the need for higher-quality data on children and their families and outlined three national goals that mirrored the point of view of children's lives and the 'whole child' perspective. These goals are listed in Table 1.2.

**Table 1.2** Three national goals of The National Children's Strategy: Our Children – Their Lives (2000)

| |
|---|
| • Children will have a voice in matters which affect them, and their views will be given due weight in accordance with their age and maturity. |
| • Children's lives will be better understood; their lives will benefit from evaluation, research and information on their needs, rights, and the effectiveness of services. |
| • Children will receive quality supports and services to promote all aspects of their development. |

Source: Department of Health and Children (2000)

Underpinned by these goals, the GUI study was designed to gain insight into the lives and development of children in Ireland and the impacts of early childhood experiences on their lives (Greene *et al.*, 2010a; Greene *et al.*, 2010b).

### 1.5.3.3 Aims and objectives

The goal of the GUI study is to gain insight into the lives of children in Ireland, in terms of children's health, education, and cognitive and socio-emotional development, set within social, economic, and cultural contexts (see Table 1.3). The GUI study was designed to longitudinally chart and study the factors that contribute to or undermine the well-being of children in contemporary Ireland. The study data and findings were intended to inform the creation of responsive, child-focused policies and improve the delivery of services for children and their families in Ireland (Greene *et al.*, 2010a).

**Table 1.3** GUI objectives

| The overarching aim of the GUI study is to inform government policy in relation to children, young people, and families.<br>To achieve this, the stated objectives of the GUI study to date are: | |
|---|---|
| **1.** | to describe the lives of children in Ireland in the relevant age categories, to establish what is typical and normal as well as what is atypical and problematic |
| **2.** | to chart the development of children over time, to examine the progress and wellbeing of children at critical periods from birth to adulthood |
| **3.** | to identify the key factors that, independently of others, most help or hinder children's development |
| **4.** | to establish the effects of early childhood experiences on later life |
| **5.** | to map dimensions of variation in children's lives |
| **6.** | to identify the persistent adverse effects that lead to social disadvantage and exclusion, educational difficulties, ill health, and deprivation |
| **7.** | to obtain children's views and opinions on their lives |
| **8.** | to provide a bank of data on the whole child |
| **9.** | to provide evidence for the creation of effective and responsive policies and services for children and families |

Source: Greene *et al.*, 2010a, p. 6.

### 1.5.3.4 Conceptual framework

The GUI study utilises a dynamic systems perspective underpinned by "five multidisciplinary insights from different disciplines: ecology, dynamic connectedness, probabilism, period effects, and the active role or agency of the child in the developmental process" (Greene *et al.*, 2010a, p. 28). These disciplines are psychology, sociology, education, public health, and epidemiology. The bioecological model devised by Urie Bronfenbrenner is core to the development of the GUI study's conceptual framework (Greene *et al.*, 2010a; Bronfenbrenner & Morris, 2006; Bronfenbrenner, 1979). Bronfenbrenner theorised that the ecological systems which define different characteristics of an environment both interact with and affect children’s growth and development. The main ecological systems are termed the microsystem (e.g. family, school and other immediate settings), mesosystem (e.g. interactions between actors in the microsystem), exosystem (e.g. local services, etc.), macrosystem (e.g. general society and culture), and chronosystem, which examines changes in major events and how the timing of these events influence one’s life (Rosa & Tudge, 2013). Elements of life course theory (Elder & Giele, 2009) were also used to shape the design of the GUI conceptual framework (Greene *et al.*, 2010a; Elder & Giele, 2009). This theory aids the understanding of how life course pathways are influenced continuously by a complex array of relationships and exposures. Multiple interacting considerations that impact on the lives of children and their life course outcomes are considered throughout the GUI study design and analysis.[5]

#### 1.5.3.5 Main study domains

The main areas of study in the GUI Cohort ’08 relate to three domains: physical health and development, social/emotional/behavioural well-being, and education/cognitive development. Examples of topics that are frequently revisited across waves include:

1. Child development, including emotional development
2. Child health, including well-being
3. Education
4. Childcare
5. Parental roles and families
6. Anthropometric measurements
7. Peer group relationships
8. Sociodemographic background

For a comparison of all selected case studies in this review by main study domains and physical and biological assessments, see Appendix 1B. See Chapter 2 (Section 2.3.1 GUI Cohort ’08) for a detailed summary of the types of data collected for GUI Cohort ’08 Waves 1 and 2.

---

[5] See Greene *et al.* (2010a, p. 24) for a graphical representation of the proposed relationships between child characteristics, child outcomes and contextual variables in the GUI study.

#### 1.5.3.6 Sampling framework

The Irish Child Benefit Register was used as a sampling frame for GUI Cohort '08 (Quail *et al.*, 2011a). This allowed for the recruitment of a nationally representative sample of infants who were aged 9 months (in their 10th month) at the time of data collection (i.e. between September 2008 and end April 2009) and their families. A total of 41,185 infants were eligible and the sample was selected on a systematic basis, stratified by parents' marital status, geolocation, nationality, and the number of children in the family unit. A random start and constant sampling fraction were utilised as a simple selection procedure. A total of 11,134 infants were recruited to the study at Wave 1. The original recruited sample was revisited at each wave, apart from expressed refusals (i.e. parents/guardians who have specifically said they no longer wish to participate), parents/guardians who have emigrated, and in cases where the cohort child has passed away. Therefore, if a cohort family has missed a wave without an expressed refusal, they are offered the opportunity to return to the study at the next wave (see Chapter 3, Section 3.3.1.2 for further details of sampling frame and recruitment).

### 1.5.4 Millennium Cohort Study

#### 1.5.4.1 Background

The MCS, also known as the Child of the New Century study, is a multidisciplinary research study that follows the lives of children in the UK (Joshi & Fitzsimons, 2016; Smith & Joshi, 2002). The study is managed by the Centre for Longitudinal Studies (CLS) in the UK. Beginning in 2000, it was the first new national birth cohort study in 30 years. Previously, there were three main birth cohort studies in Britain: the Medical Research Council's National Survey of Health and Development (from 1946); the National Child Development Study (from 1958); and the 1970 British Cohort Study (from 1970). The MCS is funded by the UK Economic and Social Research Council and a range of national Government Departments (Smith & Joshi, 2002).

The MCS follows the lives of approximately 19,000 cohort children and their families, who were recruited across England, Scotland, Wales, and Northern Ireland between 2000 and 2002. The first phase of data collection took place between 1 September 2000 and 31 August 2001 (England and Wales) and between 24 November 2000 and

11 January 2002 (Scotland and Northern Ireland). To date, there have been seven rounds of data collection for the MCS: at 9 months (2001–2002), 3 years (2003–2004), 5 years (2006), 7 years (2008), 11 years (2012), 14 years (2015) and 17 years (2019) (Connelly & Platt, 2014). MCS participants (at approximately age 19 years) also took part in three waves of online COVID-19 surveys between May 2020 and February/March 2021, alongside participants from four other national longitudinal cohort studies. These surveys aimed to provide insights into the social, health, and economic consequences of the COVID-19 pandemic. Participants were asked to provide blood samples to be tested for COVID-19 antibodies.[6] The eighth phase of data collection is expected to take place between 2022 and 2023, when the cohort participants will be around 22 years of age.

Data from the MCS have been linked via the UK Data Service to several administrative databases (e.g. birth registration and hospital attendance data, geolocation data, educational databases) (Tingay *et al.*, 2019; Hockley *et al.*, 2008). For example, National Pupil Database records of General Certificate of Secondary Education (GCSE) results have been used to link geolocation measures and educational outcomes to MCS respondents' data.

### 1.5.4.2 Rationale

The MCS was developed to chart and document the early life of the 'Children of the New Century' using a multidimensional longitudinal birth cohort design. The study collects a diverse range of information about children and their families across waves and periods of data collection to understand how early years development can influence later life (Connelly & Platt, 2014; Smith & Joshi, 2002).

### 1.5.4.3 Aims and objectives

As outlined in Table 1.4, the main aim of the MCS is to chart the lives of UK-based children and their families using key variables in the domains of health, education, social studies, and economics to investigate how the life course is impacted by a range of circumstances and conditions (Connelly & Platt, 2014; Smith & Joshi, 2002). Second, the rich range of data collected facilitates detailed analysis of how life

---

[6] See https://cls.ucl.ac.uk/covid-19-survey/ for further details.

changes and events affect children's lives over time. Third, the study focuses on parents/guardians in order to establish the children's background (Connelly & Platt, 2014). Finally, the study seeks to allow for comparison with earlier UK-based birth cohort studies. More information can be found about data integration and harmonised datasets which allow for the comparison of the MCS with other UK longitudinal studies (e.g. harmonised socioeconomic data; height, weight, and body mass index (BMI) measures). These resources were developed by CLOSER[7] and are available via the UK Data Service website.

**Table 1.4** MCS Objectives (from MCS report to funders, 2001)

| | |
|---|---|
| **1.** | To chart the initial conditions facing new children in the new century in terms of social, economic, and health advantages and disadvantages, building evidence for future research on individual development. |
| **2.** | To provide a basis for comparing processes of development with the preceding British cohorts. |
| **3.** | To collect information on previously neglected topics, such as the role of fathers, non-parental childcare, and ethnicity. |
| **4.** | To focus on the experience and aspiration of the children's parents as the immediate 'background', of the child's early years |
| **5.** | To emphasise intergenerational links including those back to the parents' own childhood. |
| **6.** | To investigate the wider social ecology of the family: social networks, civic engagement, community facilities, and services, splicing in geocoded data as available. |
| **7.** | To cover the whole of the United Kingdom, providing big enough samples for analysis within Wales, Scotland, and Northern Ireland. |
| **8.** | To provide evidence for use in the national evaluations of Sure Start and of the Children's Fund. |
| **9.** | To enhance the content of the survey by collecting data from sources beyond survey interviews, drawing on supplementary sources of funding if necessary. |

Source: Joshi & Fitzsimons, 2016, p. 410

---

[7] 'CLOSER' refers to an interdisciplinary partnership of leading social and biomedical longitudinal population studies, the UK Data Service and The British Library. See https://www.closer.ac.uk/ for further information.

### 1.5.4.4 Conceptual framework

Although not outlined in great depth in the MCS's published study documents, Bronfenbrenner's (1979) bioecological model was used to underpin the multidimensional design of the MCS (Joshi & Fitzsimons, 2016). Bronfenbrenner's model provides a framework to understand child development as a varied range of systems which children are engaged in, including family, social setting, school, and cultural systems (Bronfenbrenner, 1979). The bioecological model is also seen in the MCS survey questions based on social capital, in observations of the neighbourhood setting, and through the linking of geocoded data on statistically distinct local area settings (Joshi & Fitzsimons, 2016).

### 1.5.4.5 Main study domains

To understand the complex lives of the cohort children, the MCS places a particular focus on the domains of child development, health, education, and the socioeconomic advantages and disadvantages encountered by cohort participants (Johnson, Atkinson & Rosenberg, 2015). Topics included:

1. Child development: developmental milestones, temperament, and behaviour
2. The general health of parents
3. Psychological assessment of parents: maternal attachment, parenting beliefs, self-esteem, life satisfaction, social support, and child–parent relationship
4. Health
5. Genetic data
6. Education
7. Socioeconomic advantages and disadvantages encountered by cohort participants: neighbourhood setting, housing, employment, and income

See Appendix 1B for a comparison of all selected case studies by main study domains and physical and biological assessments. See Chapter 2 (Section 2.3.2 Millennium Cohort Study) for a detailed summary of the content of data collected for MCS Waves 1 and 2.

#### 1.5.4.6 Sampling framework

Cohort children and their families were accessed via the UK Child Benefit register. A total of 18,818 participated. The children were living in the UK at age 9 months, with a primary caregiver who was eligible to receive child benefit payments (Connelly & Platt, 2014). The sample of births was implemented over 1 year instead of 1 week in order to include births in all seasons. Participants were selected from a range of electoral wards to facilitate data analysis by neighbourhood setting. The MCS oversampled participants from areas of socioeconomic deprivation and areas with larger ethnic minority populations. Boosted samples were taken in Northern Ireland, Scotland, and Wales. Of note, Northern Ireland had not been included in previous UK-based birth cohort studies.

### 1.5.5 Danish National Birth Cohort (Denmark)

#### 1.5.5.1 Background

The DNBC study is the largest Scandinavian birth cohort study and has recruited 100,418 pregnancies of approximately 92,000 women in Denmark (Olsen *et al.*, 2001). A total of 92,670 children were born into the cohort during the period from 1996 to 2002. The DNBC was designed as a prenatal to end of life birth cohort, in that it follows children from early pregnancy until the end of life. The DNBC is a prospective data collection design with repeated measures over time. It covers a broad range of health-related domains and explores the short- and long-term effects of a range of health exposures during pregnancy and early life (Olsen, 2012; Olsen *et al.*, 2001). It was one of the first birth cohorts of its kind to start data collection during pregnancy in order to enable prenatal data collection on exposures (Olsen *et al.*, 2001).

Registry linkages play a key role in the DNBC data collection design. The DNBC data framework system links to hospital records and birth data using the unique 10-digit personal identification number (PIN) assigned to all Danish residents at birth (Olsen, 2012). Residents in Denmark with approved immigration status are also assigned a PIN. The study is linked to the national registries and data from Statistics Denmark containing medical and sociodemographic information on the whole Danish population (Schmidt *et al.*, 2019). All data from Statistics Denmark and Danish national registries are anonymised and accessed via secure Statistics Denmark servers.[8]

DNBC participants who had taken part in previous data collection waves and had provided their email address and/or telephone number were invited to take part in seven weekly COVID-19 surveys between March and June 2020 (see Chapter 3, Section 3.3.2.6 for information on the timing of additional DNBC data collection waves and response rates). A total of 53,323 adolescents (aged 16–24 years) and 53,968 mothers were invited to take part in the first wave of data collection. The surveys aimed to provide insights into how COVID-19 affected quality of life and mental health, and how participants complied with safety measures. If participants had been infected with COVID-19, data were collected on their experiences of symptoms.[9]

### 1.5.5.2 Rationale

The overall purpose of the DNBC is to establish a research database in Denmark consisting of information on exposures and other factors from conception to early life that could be linked to health-related phenomena occurring throughout life (Olsen *et al.*, 2001). The DNBC charted data to understand how maternal health and lifestyle during pregnancy affects foetal health and perinatal outcomes; specifically, how early-life exposures may impact the risk of disease across the life span (Olsen *et al.*, 2001). The DNBC investigates how prenatal and early-life exposures during foetal growth and development may link to diseases such as cancer, cardiovascular disease, asthma, and allergies (Olsen & Meder, 2014; Olsen *et al.*, 2001). To achieve this, comprehensive data on social factors, lifestyle behaviours, food intake, and environmental exposures were collected via questionnaires and registry linkages during pregnancy and when the child reached ages 6 months, 18 months, and 7 years (see Chapter 2, Section 2.3.3 for more details on DNBC data collection and administrative database linkage).

---

[8] See https://www.dst.dk/en/TilSalg/Forskningsservice for further details about Statistics Denmark and its privacy policies.

[9] For more information about the COVID-19 surveys, see https://www.dnbc.dk/data-available/covid-19.

### 1.5.5.3 Aims and objectives

The main study aims and objectives (Olsen & Meder, 2014; Olsen *et al.*, 2001) are to:

- Study pregnancy complications and disease in children to understand the factors and short- and long-term consequences of exposures that affect foetal development and early life health.
- Study diseases in children which are thought to originate during the foetal development period.
- Determine the side-effects of medications and infections on children during the prenatal and early childhood period.
- Collect data during pregnancy to create an exposure register that could link with active disease registers so that prenatally collected data could be studied.
- Create a medicines registry of all medications used during pregnancy to establish whether there are any prenatal side-effects associated with certain medications.
- Develop a Danish research information bank to aid future research.

### 1.5.5.4 Conceptual framework

The DNBC study is underpinned by the Fetal Origins Hypothesis, also known as the Barker Early Origins Hypothesis (Barker, 1995). David Barker first proposed this concept in 1986 (Barker & Osmond, 1986) and went on to develop the hypothesis further in the late 1980s and early 1990s. Initially, his research made a direct link between the effects of in utero nutrition on late-onset chronic heart disease (Almond & Currie, 2011; Barker, 1995). Barker found that low birthweight and its possible associated effects on organ development and function could be linked to adverse epidemiological well-being in later life (Barker, 1995). For example, Barker contended that low weight in utero can result in overweight adults, leading to diseases associated with obesity such as cardiovascular disease, diabetes, and hypertension (Barker, 1995). Barker argued that the under-nourished foetus prioritises vital organ development at the expense of the ideal development of other organs, and that this less optimal organ development may be related to the development of diseases throughout life (Barker, 2007; Barker, 1995). Decades later, research has shown that exposures during prenatal development can impact on a wide range of life-long health issues (Gluckman *et al.*, 2008; Gluckman & Hanson, 2004). While much of the primary DNBC study literature references Barker's work, the Fetal Origins Hypothesis

(Barker, 1995) is not explicitly defined as the conceptual framework for the study. Indeed, no information on the study's conceptual framework can be sourced. Yet, what is evident is that the mechanisms underlying Barker's hypothesis aided the development of the DNBC design with regard to the impact that exposures in the womb can have on health in later life (Olsen & Meder, 2014; Olsen *et al.*, 2001).

### 1.5.5.5 Main study domains

Study domains included health data, alongside related information regarding lifestyle, socioeconomic factors, diet, and mental well-being (Nybo Andersen & Olsen, 2011). Topics included:

1. Birth outcomes
2. Health data: obesity and physical exercise during pregnancy; infections and medication during pregnancy; environmental toxins
3. Exposures and related information regarding lifestyle: alcohol, coffee, and nicotine use during pregnancy
4. Mental well-being
5. Maternal working conditions and education
6. Diet
7. Biological samples (child and mother).

For follow-up waves of the study, data collection centred around health and developmental outcomes (e.g. inoculations and vaccinations, medicine, diet and exercise, height and weight, social development, mental health and body image, dental health, mobile phone use, education, puberty, and the impact of COVID-19). For further details, see the DNBC study website.[10]

### 1.5.5.6 Sampling framework

Candidates for the cohort were all pregnant women living in Denmark between 1995 and 2002. The women were ideally recruited as early as possible in their pregnancy (between 6 and 12 weeks) but could be recruited up until 24 weeks of pregnancy (Nybo Andersen & Olsen, 2011). Recruitment was conducted in collaboration with

---

[10] https://www.dnbc.dk/

general practitioners (GPs), who played an essential role in the first antenatal visit. A total of 100,418 pregnant women were recruited into the study and 92,670 children were born into the cohort. Less than 1% of participants had formally withdrawn from the study by the time of the 7-year follow-up, and most participants have engaged in the new data collection rounds (Greene *et al.*, 2011) (see Chapter 3, Section 3.3.2 for comprehensive details on data collection rounds). There have been several data sweeps throughout childhood and adulthood (at 7 years, 11 years, 18 years, and 16–24 years), and with the mothers of the study children.

## 1.5.6 Étude Longitudinale Française depuis l'Enfance (France)

### 1.5.6.1 Background

The ELFE study is the first national longitudinal birth cohort study of its kind in France (Charles *et al.*, 2020). The study cohort includes 18,329 infants born in France in 2011, studied from birth to age 20 years. Prior to the ELFE study, there were several regional birth cohort studies of children to research the impact of exposure to environmental contaminants on child health outcomes (Charles *et al.*, 2020). In the 2000s, two national projects were created. The first project, led by the French Institute for Demographic Studies (L'Institut national d'études démographiques; INED), studied children to better understand their life circumstances, with an emphasis on socioeconomic and health inequalities (Pirus *et al.*, 2010). The second project was run by the French Institute for Public Health Surveillance (Institut de veille sanitaire; InVS) as part of the first French National Environmental Health Plan to collect national data regarding the level of pollutant exposures experienced by pregnant women and by children (Pirus *et al.*, 2010). The two projects were merged into the ELFE study and data collection officially began in 2011.

Data collection in the first 5 years consisted of telephone interviews of both parents when the child was aged 2 months, 1 year, and 2 years, and of one parent when the child was aged 3.5 years; and a home visit when the child reached age 3.5 years (Charles *et al.*, 2020). To gain further insights into the children's development, contact was also made with teachers, school medical officers, GPs, and infant welfare services. ELFE data were linked with data from various health insurance schemes operated through France's social security data management system (Système National d'Informations Inter-Régimes de l'Assurance Maladie; SNIIRAM) to complement the health information provided by parents, particularly about any prescribed medication (Charles *et al.*, 2011).

ELFE study participants also took part in two COVID-19 surveys between April and May 2020. The cohort participants' parents were surveyed to explore the impact of COVID-19 on family life and the health and behaviour of their child(ren) during this time. Data collected related to child health during the COVID-19 pandemic, home schooling, family tactics to prevent possible infection, coping strategies, and child mental well-being.[11] Some participants were also asked to provide serological samples to be tested for COVID-19 antibodies. France's National Institute of Health and Medical Research (Institut national de la santé et de la recherche médicale; INSERM) and INED requested for these data to be collected by ELFE and by other epidemiological studies in France to inform relevant public health prevention policies.[12]

### 1.5.6.2 Rationale

Although there were several area-specific birth cohort studies already in progress in France, the ELFE study was devised to understand the pre-birth and early-life influences, and to study the attribution of a range of factors that influence health, development, and well-being (Vandentorren *et al.,* 2009). The rationale for this cohort study stems from the awareness that, from conception to the young adult years, individuals move through different stages of growth and development, and that prenatal and early-life exposures (i.e. nutrition, pollution) within individuals' environments can affect their life-long health and create a range of health issues (Vandentorren *et al.*, 2009).

### 1.5.6.3 Aims and objectives

The primary aim of the ELFE study is to examine the contributors to child health, growth, development, and social skills from birth to adulthood using a

---

[11] For further information on the ELFE COVID-19 survey, see https://www.elfe-france.fr/en/teachers-section/covid-19-survey/.

[12] For further information on the ELFE COVID-19 survey rationale, see https://www.elfe-france.fr/en/teachers-section/covid-19-survey/.

multidisciplinary approach (i.e. epidemiology, public health, and social sciences) (Charles *et al.*, 2020; Vandentorren *et al.*, 2009). Other specific study objectives are to:

- Assess the impact of environmental pollutants during pregnancy and early life on a child's neurocognitive and reproductive development.
- Understand the impact of family characteristics (i.e. blended families, separation, divorce) and socioeconomic factors on families and their engagement with the French education system.

Study childhood development and how a range of factors interact throughout a child's life and impact on their social and educational pathways. These factors include family setting, school setting, health, and diet.

### 1.5.6.4 Conceptual framework

The conceptual framework utilised by ELFE is a life course approach to chronic disease (Kuh *et al.*, 2003; Ben-Shlomo & Kuh, 2002). This framework is used to understand how chronic disease development is affected by genetic, behavioural, and psychosocial exposures during the prenatal period through to adulthood, and how factors within these exposures interact over time. This conceptual framework was selected for ELFE to study and monitor child development and provide researchers with data to understand the interaction between different settings, life paths, and life course events and their impact on children's lives (Charles *et al.*, 2020; Pirus *et al.*, 2010). In particular, this conceptual framework underpins the study of the long-term risk of chronic disease posed by environmental pollutant exposures, and other physical and social exposures during pregnancy and early life. The life course approach to chronic disease framework was also considered to be useful for the study in the context of the broader areas of health and social trajectories (Charles *et al.*, 2020; Pirus *et al.*, 2010).

### 1.5.6.5 Main study domains

The ELFE study covers a range of study domains on child development and health, with a specific focus on environmental exposures (Vandentorren *et al.*, 2009). Topics include:

1. Child development and health

2. Environmental exposure and health
3. Social and nutritional aspects of the child's feeding and their relationship with development and health
4. Development of updated growth reference curves
5. Child cognitive development
6. Asthma and allergies incidence and early determinants
7. Family environment, events, and social outcomes, including children's living conditions, generational relations and first experiences, and the construction of children's sexual identities
8. Parental health, behaviour, and life values
9. Educational outcomes

#### 1.5.6.6 Sampling framework

A total of 18,329 children were recruited at birth in a random sample of 349 maternity hospitals selected from the 544 public and private maternity hospitals in metropolitan France. Recruitment took place over 25 days during four set periods in 2011. The inclusion criteria were women aged over 18 years who gave birth after at least 33 weeks' gestation to single or twin children, and who did not plan to leave metropolitan France within 3 years of giving birth. Just over 50% of new mothers who were informed about the study agreed to participate (Charles *et al.*, 2020; Vandentorren *et al.*, 2009). To ensure representativeness, families who did not hold French citizenship but were living in France could also participate, provided that the birth mother could read French, Arabic, Turkish, or English.

### 1.5.7 Generation Victoria (Australia)

GenV is the newest of the birth cohort case studies included in this report. The first phase of recruitment for GenV began in mid-2021 and will be ongoing until 2023. Therefore, this case may not be discussed as much as the other cohort studies in later chapters. For example, as data collection has just commenced, data analysis methods and dissemination strategies will not be discussed in detail in later chapters. Nevertheless, GenV was selected for inclusion for several important reasons. First, this is a new study using state-of-the-art methodologies to develop a large-scale, parallel (i.e. children and parents), Victoria-wide birth cohort.[13] The study will collect a rich range of data through biobanks, health and education data linkage, and through low-burden data collection instruments (i.e. surveys, videos, and games ranging from approximately 3 minutes to 20 minutes). Any learnings regarding minimal burden on participants and retention are likely to be of great value to DCEDIY. Second, the GenV conceptual framework has been guided by a detailed and clearly described selection and development process. This work has been published and is a valuable tool for birth cohort methodologists and researchers in the design of new birth cohort studies (Wang *et al.*, 2021).

### 1.5.7.1 Background

While there is a history of longitudinal studies relating to health and well-being in Australia (Christensen *et al.*, 2017), GenV is the largest ever longitudinal birth cohort to be undertaken in Australia (Wake *et al.*, 2020). The study uses a cutting-edge research design to enhance childhood research and policy within the Australian context, with acknowledgement that the Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC) already exists (Wake *et al.*, 2020). LSAC showed that from the age of 2 years most children were suffering from at least one health issue, and that health issues were likely to persist consistently if prevailing for more than 2 years (Liu *et al.*, 2018). GenV was designed to explore and chart the health and social impact of a range of factors across childhood into adulthood (Wang *et al.*, 2021). Another primary goal of the study is to reduce the time and expense that birth cohort studies are widely known for, and to reduce participant burden through the additional use of state-of-the-art biobanks and retrospective and prospective linkage to clinical and administrative datasets (Wake *et al.*, 2020). Participant data collection waves will roll out approximately four times per year over the first years of the children’s lives.

### 1.5.7.2 Rationale

The purpose of GenV is to undertake a large-scale, state-wide birth cohort study to understand and address complex health and well-being issues affecting infants into

---

[13] For more information on the GenV study, see https://genv.org.au/.

adulthood. The study focuses on the state of Victoria, with the aim of generating translatable research in the areas of health forecasting, prevention, service provision, and disease treatment (Davies *et al.*, 2020).

### 1.5.7.3 Aims and objectives

GenV aims to explore causal pathways between environmental exposures, genetics, and physical and mental health to chart their health and social impacts across childhood and into adulthood (Wang *et al.*, 2021). The primary objective of GenV is to create a large-scale, parallel birth and parent cohort study in the state of Victoria for intervention and prevention research (Wake *et al.*, 2020). The study also intends to blend collected data with linked data and biosamples to build upon observational health services and policy research.

### 1.5.7.4 Conceptual framework

The GenV study is underpinned by two conceptual frameworks. The primary framework utilised is Shonkoff's biodevelopmental framework (Shonkoff, 2010). The framework was developed to enable policy-makers to study the causal pathways (e.g. distorted gene expression) through which environments affect child development, physical and mental health, and learning (Shonkoff, 2010). The framework is built upon Bronfenbrenner's ecological theory (Bronfenbrenner, 1994). The biodevelopmental framework offers a cohesive human development theory and knowledge-based framework that can support integrated improvements to promote childhood health and development and reduce inequalities (Wang *et al.*, 2021; Shonkoff, 2010). The framework's integrated approach and the clarity of its visual summary were additional reasons for its selection as the primary conceptual framework for the GenV study (Wang *et al.*, 2021).

GenV utilises a secondary conceptual framework to underpin and support the study. The AIHW's (2021) person-centred model was selected to better understand the social determinants of health and well-being within an Australian setting (Wang *et al.*, 2021). The model reports across seven different domains, including health, household income, and family social support.

#### 1.5.7.5 Main study domains

GenV uses a six-domain 'focus area framework' to guide the study's focus and data collection around the overarching theme of childhood inequalities. The six domains are: (1) healthy environments, (2) development and learning, (3) obesity and diabetes, (4) mental health and well-being, (5) organ health, and (6) infection, immunity, and allergies.[14]

GenV will also use repeat principal health and well-being measures comprising health-related quality of life and diagnostic estimates of disease/disability burden (Sung *et al.*, 2021). Additional measures will combine service-related data (e.g. costs, medications) for economic analyses. GenV plans to collect biosamples of multiple tissues (e.g. blood, saliva, stool, breast milk) at several time points, including all pregnancy trimesters, after birth, and at school entry. See Appendix 1B for a comparison of all selected case studies by main study domains and physical and biological assessments.

#### 1.5.7.6 Sampling framework

The sampling frame is open to all babies born in Victoria between mid-2021 and mid-2023, and their parents. Approximately 75,000–80,000 babies are born each year in the state's hospitals (Wake *et al.*, 2020). It was estimated that by the end of 2022, the parents/guardians of more than 170,000 babies would have been invited to participate in the study. To be included in the study, parents/guardians must have the decisional capability to consent to participate in GenV. To support participant retention, if parents and children are relocating from Victoria at any time, they may remain in the study via linked and contributed data (Wake *et al.*, 2020).

### 1.5.8 The Early Life Cohort Feasibility Study (UK)

Although not a named case study for the purpose of this review, the ELC-FS[15] is a new UK-based feasibility study which will offer valuable insights into a range of learnings that will benefit future birth cohort study research teams. The ELC-FS was

---

[14] For a graphical depiction of GenV's focus area framework, see https://www.mcri.edu.au/images/mcri/partnerships/our-supporters/Genv-Focus-Areas-Diagram.pdf.

[15] For additional information on the ELC-FS, see https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study/.

launched in April 2021 as a feasibility study to test new methodologies and approaches to implementing a large-scale birth cohort study. For example, between June and September 2021, the ELC-FS team consulted widely with researchers, data users, and policy-makers regarding the planned content and design of the proposed study. The study also aims to collect a range of robust data from a representative cohort of infants to uncover the challenges facing the current generation at such a vital period in the UK's history (i.e. post-Brexit, and during the COVID-19 pandemic).

The primary objectives of the ELC-FS are to:

- Test the feasibility of accessing and maintaining a comprehensive recruitment strategy and sampling frame of babies born in England, Wales, Scotland, and Northern Ireland. The goal is to develop a robust sampling framework which will focus on the overall retention of participants, while also maximising the engagement and retention of peoples from hard-to-reach groups and deprived areas.
- Collect data on cohort children's health, development, well-being, and socioeconomic settings.
- Use innovative techniques and novel measures to collect data on parent–child interaction and children's function.
- Collect biosamples from cohort children and their parents.
- Develop a strategy to link the study data to electronic administrative data and geo-environmental data which will be rolled out for the main study, if commissioned.
- Engage members of the public, alongside policy-makers, practitioner groups and networks, and researchers who work with data to determine needs and priorities for the study.

The study is funded by the UK Economic and Social Research Council (ESRC) until March 2023. An evaluation phase will then commence close to the end of the feasibility study to determine whether a main study will be launched. At the time of publication of this review, there are no technical documents or primary findings published regarding this study.

### 1.5.9 Growing Up In Digital Europe

GUIDE will be Europe's first cross-country comparative birth cohort study. The main objective of GUIDE is to provide high-quality longitudinal data to support policies for improving the well-being of children, young people, and their families across Europe. The core concept of well-being will inform the study, incorporating both subjective (i.e. positive affect, negative affect, and overall life satisfaction) and psychological aspects. Using an accelerated longitudinal design, GUIDE will follow a sample of newborn infants and a sample of school-age children. The study is currently in its design and preparation phase (2022–2026). The older cohort will first be studied at approximately age 8–9 years, and subsequently followed up every 3 years until the age of 23–24 years. The younger cohort will first be studied at age 0–1 years at Wave 1, followed up at age 2 years, and then revisited every 3 years until the age of 23–24 years in 2052. The survey is currently planned to take place in 22 countries, with plans to extend this to all countries in Europe. All countries will use the same questionnaires, translated into local languages. The best possible random sampling practice will be used in each country, even if the sample design differs across countries due to national variation in population density and distribution of schools and residences (Lynn, 2019). Sampling frames will provide the best possible coverage of target populations. Pilot surveys are being carried out in a subset of countries (France, Croatia, and Finland) throughout 2022–2023. Questionnaire tools will be tested first to ensure translation accuracy. The first wave of fieldwork is anticipated for 2027.

## 1.6 Strengths and limitations of the case study conceptual frameworks

A key strength of the GUI and MCS conceptual frameworks is the facilitation of multidisciplinary analyses of socioeconomic inequalities across a range of child outcomes over time. The collection of data on an array of parental and family characteristics, as well as those of schools and neighbourhoods, enables researchers to ascertain the influence of these socialising agents on child developmental trajectories. This is beneficial to social scientists and policy-makers interested in the evolution of socioeconomic inequalities (e.g. socioeconomic disparities in child cognitive ability from preschool age through compulsory schooling) (Skopek & Passaretta, 2021). Moreover, the conceptual framework underpinning the GUI study was successfully extended to adulthood; the conceptual framework for Cohort ’98 was adapted at Wave 3 (age 17/18 years) to incorporate other relevant theoretical perspectives for understanding the transition into young adulthood (McNamara *et al.*, 2020).

In contrast, the DNBC sought to examine the links between pregnancy conditions and the child's intrauterine environment, and subsequent health outcomes in infancy, childhood, and adulthood. Although the technical study documentation does not provide much detail on the conceptual underpinnings of the DNBC, its objectives are consistent with the Fetal Origins Hypothesis, which outlines the foetal and infant origins of adult disease (Barker, 2001). DNBC Principal Investigator Olsen (2012) argues that the DNBC's large-scale nature, long-term follow-up period, low attrition rates and inclusion of biological samples makes this cohort particularly well-suited to testing hypotheses regarding the links between early exposures and adult health outcomes, including the role of genetics and epigenetics. Thus, a key strength of the DNBC conceptual framework is that it focuses on the whole life course. Although the GUI study has adapted its conceptual framework to embrace a life course approach and is currently planning fieldwork for Cohort '98 at age 25 years (McNamara *et al.*, 2020), additional study waves into adulthood have not yet been confirmed. Conversely, the DNBC is designed to follow children into adulthood and old age. In fact, the DNBC team dubbed it a "conception-to-death cohort" (Nybo Andersen & Olsen, 2011, p. 115). However, compared to the studies rooted in Bronfenbrenner's bioecological model, the DNBC collects a more limited range of contextual data on the various environments (e.g. family, school, neighbourhood) participants are embedded in. While the DNBC's linkages to administrative data cover some non-medical outcomes over the life course, such as educational achievement (e.g. Keilow *et al.*, 2019), parental and child interviews focus predominantly on health-related measures rather than socioeconomic ones. Although socially determined health inequalities may be linked with employment (hence derived social class) and income data for participants, the DNBC is less suitable for robust secondary analyses of socioeconomic inequalities across the life course.

Meanwhile, ELFE draws on Kuh and Ben-Shlomo's (2004) life course approach to chronic disease epidemiology. The ELFE team chose this framework to inform their longitudinal study because it emphasises the relevance of critical early life periods for children's health outcomes via familial, social, and cultural pathways (Charles *et al.*, 2020). According to Wang *et al.* (2021), the key advantages of Kuh and Ben-Shlomo's approach in informing longitudinal cohort studies is that it considers the intergenerational transmission of inequalities via genetic and/or social pathways, incorporating several levels of influence (e.g. household, neighbourhood, and national) both across individuals and over time. The main weakness of the framework is that it focuses primarily on chronic disease. However, ELFE's conceptual framework is broader in scope than Kuh and Ben-Shlomo's original model. Moreover, unlike the studies underpinned by Bronfenbrenner's bioecological model, ELFE's conceptual framework is relevant to adult outcomes.

GenV relies on a purposefully chosen conceptual framework. Wang *et al.* (2021) described the selection process for the conceptual framework: a pragmatic literature search yielded 14 life course frameworks which were assessed against 7 criteria informed by the principles of the Life Course Health Development Framework (Halfon & Forrest, 2018). The framework had to: (1) be broad in scope; (2) be multidimensional and interdisciplinary; (3) encompass physical, mental, and social well-being outcomes; (4) include individual life span and intergenerational aspects; (5) allow multilevel and interacting pathways of influence; (6) be multi-age; and (7) be embedded in a visually clear infographic. Shonkoff's (2010) biodevelopmental framework was chosen as the primary conceptual framework to underpin GenV as it met all of the outlined criteria. The framework's original conceptualisation focused on early childhood, so it had to be adapted to include potential feedback loops across the life course and intergenerational effects (Wang *et al.*, 2021).

To conclude, large-scale longitudinal studies often rely on conceptual frameworks relevant to their aims, objectives, and disciplinary focus in order to guide the study design, scope, and measures, alongside the eventual primary and secondary analyses. A review of the conceptual underpinnings of the case studies summarised in this report suggests two main takeaways. First, there is variation in the extent to which study documentation makes the choice of a framework explicit. GenV offers a best-practice scenario because its documentation not only explains how the study's conceptual underpinnings inform its design (as per the GUI study documentation), but also elucidates the systematic process used to identify the specific framework chosen. Second, existing life course theories do not always offer a comprehensive guide to a new study but leave room for adaptation. For example, the new UK-wide ELC-FS is similar to the MCS and the GUI study in its focus on the evolution of inequalities in child development. However, in addition to the home, environmental, and neighbourhood influences on these trajectories, the ELC-FS also aims to investigate their biological determinants. This is, of course, entirely in line with Bronfenbrenner’s bioecological model, which (unlike its precursor, ecological systems theory) extends to biological determinants and their interactions within the social context. Meanwhile, GenV used Shonkoff’s model as a starting point and then incorporated further life span and intergenerational elements to align this framework with the study objectives.

## 1.7 Conclusion

This chapter introduced four longitudinal birth cohort studies that will be discussed throughout this report: the MCS, DNBC, ELFE, and GenV studies. These case studies were selected because they are large-scale, recent, nationally or regionally representative, broad in their scope and coverage, and innovative in their methods of data collection. The studies’ rationales, objectives, conceptual frameworks, key topic domains, and sampling strategies were briefly summarised alongside those of the foundation case study used in this report: the GUI Cohort ’08. Of the four comparator case studies, the MCS is the most similar to GUI ’08, while the DNBC and the new GenV study are most dissimilar. Although GenV is the newest of the four (having launched recruitment in mid-2021), it is one of the most cutting-edge longitudinal birth cohort studies to date, offering multiple lessons to learn. Chapter 2 will review the main domains and modes of data collection used in the case studies.

# Appendices

## Appendix 1A

**Table 1.5** Comparison of selected birth cohort case studies by study features

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) | GenV (Australia) |
|---|---|---|---|---|---|
| Starting year | 2008 | 2000 | 1996 | 2011 | 2021 |
| Age at entry | 9 months | 9 months | Pre-birth | Pre-birth | From birth |
| Typology | Multidimensional, interdisciplinary, multi-ministerial longitudinal survey | Multidisciplinary study | Specialised health-focused study | Specialised health-focused study | Specialised health-focused study |
| Conceptual framework | Bronfenbrenner's bioecological model | Bronfenbrenner's bioecological model | Barker's Fetal Origins Hypothesis | Life course approach to chronic disease conceptual framework | Shonkoff's biodevelopmental framework<br>The AIHW's person-centred model |
| Focus | Multidisciplinary with emphasis on policy | Multidisciplinary | Health, including reproductive issues experienced by pregnant women | Multidisciplinary, specific environmental focus regarding health | Multidisciplinary, specific focus regarding health |
| Initial sample size | ≈11,134 | ≈19,000 | ≈ 92,892 | ≈20,000 | ≈170,000 (expected) |

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) | GenV (Australia) |
|---|---|---|---|---|---|
| Sampling design | Systematic basis, pre-stratifying by marital status, county of residence, nationality of payee, and number of children on child benefit | Geographically clustered, with ethnic and disadvantage oversampling | Women at their first antenatal visit to GP | All children born over the course of 16 days (during the same year) in 344 selected maternity units | Open to all babies born in birthing hospitals in the state of Victoria over 2 consecutive years |
| Sampling frame | Child benefit records | Child benefit records | Pregnant women living in Denmark with Danish PINs and their child(ren) | Not applicable (see above) | Not applicable (see above) |
| Data collection phases | 9 months, 3 years, 5 years, 7/8 years, 9 years, 12/13 years | 9 months, 3 years, 5 years, 11 years, 14 years, 17 years, 22 years | Prenatal, 6 months, 18 months, 7 years, 11 years, follow up among mothers, dietary habits of 14-year-olds, puberty follow-up, 18 years, COVID-19 follow-up (16-24 years) | Prenatal, 2 months, 3–10 months, 1 year, 2 years, 3.5 years, 4 years, 7 years, 8 years, 10.5 years | Birth, every year after birth for 4 years, school health check |
| Data linkages | National Perinatal Reporting System (NPRS) data (see Thornton *et al.*, 2013) | Health system, school, census, childcare, maternity, and birth | Health system, maternity and birth, social and occupational status (registries) | Health system, school, census, childcare, environment, social security, maternity, and birth | Government Departments, hospitals, doctors, and other providers relating to participants' health, mental health, education, and social records and services |

# Appendix 1B

**Table 1.6** Comparison of selected birth cohort case studies by main study domains and physical and biological assessments (topics in each domain are indicative)

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) | GenV (Australia) |
|---|---|---|---|---|---|
| **Main study domains** | **Pregnancy and childbirth**<br>**Social/emotional development/child behaviour**<br>Family environment, play/leisure time activities, emotional development, parental role, parental participation, peer relationships, screen time, mental health<br>**Physical health and development**<br>Physical health, general health, physiological development, chronic conditions and disabilities, special education needs, health behaviours, anthropometric data | **Pregnancy and childbirth**<br>**Child development and behaviour**<br>Direct assessments of cognitive and physical development. Parents' report of child's behaviour such as strengths and difficulties questionnaire and questions about socio-emotional development and child's day-to-day activities.<br>**Education and childcare**<br>Experiences at school, details of school attended, school choice, current use of childcare and childcare history<br>**Child health**<br>General health, eyesight, hearing, dental health, immunisations, accidents, diagnosed conditions and medication, diet and physical activity | **Birth outcomes**<br>**Health data**<br>Obesity and physical exercise during pregnancy; infections and medication during pregnancy; environmental toxin exposures and related information regarding lifestyle: alcohol, coffee, and nicotine use during pregnancy<br>**Biosamples**<br>**Mental well-being**<br>**Maternal working conditions and education**<br>**Diet** | **Child development and health**<br>Social and nutritional aspects of child's feeding and their relationship with development and health, development of updated growth reference curves, child cognitive development, asthma and allergies incidence and early determinants<br>**Childhood environmental exposures**<br>Environmental exposure and health<br>**Education and childcare**<br>Educational outcomes, school behaviour and | **Health data**<br>Healthy environments, organ health, obesity and diabetes, infection, immunity and allergy, health-related quality of life and diagnostic estimates of disease/disability burden<br>**Biosamples**<br>**Development and learning**<br>**Mental health and well-being** |

| | GUI Cohort ’08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) | GenV (Australia) |
|---|---|---|---|---|---|
| | **Education/cognitive processes**<br>Mental development, child day care, educational environment, teacher–student relationship, educational engagement, achievement, aspirations, civic and economic engagement<br>**Parenting**<br>**Parental employment, education, and income**<br>**Sociodemographics**<br>Occupation/social class, household income, family structure/type, parental educational attainment, data on neighbourhood context | **Family demographics**<br>Changes in family composition and wider family demographics, family relationships – parents (resident and non-resident) and child, child and siblings, parents with each other<br>**Parenting**<br>Parenting behaviours and attitudes<br>**Parental health**<br>Physical health (including height and weight) and health behaviours, plus mental health<br>**Parental employment, education, and income**<br>Including employment histories and work–life balance<br>**Social and community context**<br>Location and neighbourhood, social support and participation, ethnicity, and language | | socialisation, day care and school<br>**Parental health, behaviour, and life values**<br>**Family life**<br>Family environment, events, and social outcomes including children’s living conditions, and generational relations<br>**Sociodemographic characteristics** | |

| | GUI Cohort ’08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) | GenV (Australia) |
|---|---|---|---|---|---|
| **Physical and biological assessments (timing: various)** | **Measured by interviewer:**<br>Head circumference, length, weight, height, waist circumference | **By interviewer/respondent:**<br>Weight, height, body fat, head circumference, waist circumference, physical activity monitoring (accelerometery), oral fluid samples, shed milk teeth | Maternal blood samples taken by GP during routine check-ups conducted at the 6th, 12th, and 24th weeks of pregnancy<br>Whole blood samples mixed with ethylenediaminetetraacetic acid (EDTA)<br>Umbilical cord blood samples taken immediately after birth at the hospital by a midwife or nurse | **By medical professionals:**<br>Height, weight, urine and blood samples (subsample), mother’s blood and urine, umbilical cord sample, mother’s hair, breast milk, baby’s first stool, home dust (subsample) | Pregnancy samples including maternal serum screening (MSS) bloods, non-invasive prenatal testing (NIPT) bloods, both collected at 10–12 weeks’ gestation, and newborn blood spot screening (NBS) collected at birth<br>Biosamples collected as part of routine care in pregnancy but discarded after testing, including pregnancy bloods taken at 10–12, 24–28, and 36 weeks’ gestation, Group B streptococcus swab collected around 34–38 weeks’ gestation<br>Saliva to be collected during participant recruitment from babies and parents if consented to. Placenta, cord blood, and neonatal stool samples may be collected. |

# Chapter 2: Data collection considerations

## 2.1 Chapter overview

Chapter 2 will begin with a narrative synthesis of the academic literature regarding international advances in the data collection methods of birth cohort studies over the past decade, including using technology for data collection and biomedical data collection, and implementing data linkage for administrative or population-based records. The feasibility of data linkage in Ireland will also be considered. Following this, case studies will focus on the domains of data collection, indicative data collection instruments across domain topics (e.g. measures, standardised scales, exemplar questions), and modes of data collection in the named cohort studies (first and second waves only). The collection of any biomedical data and/or linkage to administrative databases (across all waves) in each of the named birth cohort studies will be reviewed.

## 2.2 Narrative synthesis of the academic literature

Birth cohort study designs may reduce research participant burden through implementing novel methods of data collection (e.g. web-based questionnaires, linkage with administrative databases) (Shamsipour *et al.*, 2020). The aim of this review is to give an overview of contemporary innovations in data collection modes and techniques, including utilising technology (e.g. online data collection, apps) for data collection, and incorporating biomedical data in cohort studies. This will be followed by a discussion of the potential for data linkage in the Republic of Ireland.

### 2.2.1 Technological advances

Technological change has significantly transformed everyday life since the turn of the millennium, and innovations in technology present valuable opportunities for collecting data at scale (Calderwood *et al.*, 2019). Originating with the use of computer-assisted self-interviewing (CASI) for inquiring about sensitive subjects (e.g. in the MCS and GUI), Internet-based platforms are progressively being used in research, as traditional modes (e.g. postal surveys, face-to-face interviews, telephone interviews) are witnessing a decline in response rates and a rise in costs (Blumenberg *et al.*, 2018). This is especially pertinent for longitudinal birth cohort studies, which are prohibitively costly to conduct on a large-scale basis (Pugh *et al.*, 2015). Nevertheless, unlike online data collection methods, face-to-face data collection enables the interviewer to directly establish a good rapport with participants and encourage long-term engagement (Thornton *et al.*, 2013). Moreover, certain measurements (e.g. cognitive tests, non-invasive biomedical sampling) can only be carried out by a trained interviewer within the home setting.

### 2.2.2 Web-based questionnaires

Considering the barriers associated with traditional modes of data collection, the use of web-based surveys is growing globally. Initial uncertainty about the potential of digital technologies for administering surveys has subsided as Internet access has become more widespread (Tienda & Koffman, 2021). As of 2019, 91% of Irish households had an Internet connection; fixed broadband was the most common means of Internet access in the household (84% versus 47% using mobile broadband), and some households used multiple types of Internet connection (CSO, 2019). Online surveys are relatively inexpensive to create and maintain, and often yield higher quality data (if filtering questions and consistency checks are employed) (Blumenberg *et al.*, 2018). Moreover, online data collection saves time by gathering data in a format suitable for analysis (Firestone *et al.*, 2015). Online data collection is also proposed in order to reduce environmental costs and accelerate the questionnaire administration process (Bray *et al.*, 2017). For instance, completing a survey online has been reported to take about half the time required to answer the same questionnaire via telephone interview (van Gelder *et al.*, 2020). Study participants are afforded greater choice regarding how and when to complete the survey, alongside instant and simple submission of data (this method reduces the need for postal reminders, locating a postbox, etc.) (Bray *et al.*, 2017).

Nevertheless, a corollary of the efficiency and affordability of web-based surveys is the absence of an interviewer to facilitate data collection and the reliance on data generated solely by participants. These data are determined by what researchers hope are high-quality, validated questionnaires (Pugh *et al.*, 2015). The validity of web-based questionnaires could also be hindered by participants' individual patterns of responding (e.g. questionnaire breakoff, item non-response) (Blumenberg *et al.*, 2018). Furthermore, it is critical to identify what leads participants to return to a web-based survey and to acknowledge the biases that may be introduced by the sociodemographic differences between those who return and those who are lost at follow-up (Loxton *et al.*, 2019).

Concerns about switching to an online-only approach for data collection are further compounded by evidence from both market research and health-related research that web-based surveys yield lower response rates compared to traditional paper and telephone modes (Bray *et al.*, 2017). Nowadays, even a response rate below 10% is not considered unusual for an online survey (Van Mol, 2017). Web-based surveys, which are typically accessed through a link that leads participants to the survey web page, can be perceived as more impersonal compared to telephone and face-to-face modes (Goodman *et al.*, 2020). Moreover, it has been proposed that the proliferation of web-based surveys in circulation inflicts an increasing burden on the public, thus reducing response rates. Associated with the growing demand for participation is the concept of survey fatigue stemming from continual invitations to provide feedback on products and services. Consequently, potential participants might find it complicated to differentiate birth cohort studies from market research (Harrison *et al.*, 2020). Reduced response rates arising from online data collection runs the risk of obtaining unrepresentative samples, which could hinder external validity and introduce bias in the estimates derived from the collected data (Harrison *et al.*, 2020).

Further evidence from a randomised controlled trial conducted with participants from the Avon Longitudinal Study of Parents and Children (ALSPAC) highlights the possible shortcomings of moving towards an online-only approach for survey administration (Bray *et al.*, 2017). This randomised controlled trial found that response rates for those offered a choice between completion methods (paper or online) were much higher than the group offered an online-only option at the outset of the study. However, offering a choice of method was more expensive, costing an average of £0.71 more per participant than offering online completion only (Bray *et al.*, 2017). Accordingly, developers of future birth cohort studies should weigh up the additional costs of offering a choice of completion method against potential improvements in response rates among those not yet prepared to move to an online-only model.

Notwithstanding this finding, large-scale birth cohorts are increasingly digitising their methods of data collection (Nkyekyer *et al.*, 2021). The collaborative international mother–child cohorts of the Nascita e INFanzia: gli Effetti dell'Ambiente (NINFEA) study (Italy) and the Early Life Factors (ELF) study (New Zealand) were pioneering in utilising the Internet for the recruitment and follow-up of their members. Launched in 2005 as a pilot study in the Italian city of Turin, the NINFEA study has been increasingly expanded to all of Italy. As of March 2015, a total of 7,003 pregnant women had registered to participate in the NINFEA cohort via the project website and had completed the initial online questionnaire. In conjunction with email, telephone calls and SMS text reminders, women accessed the various follow-up questionnaires through the study website using their username and password. Of the total pregnant women recruited at baseline, 88% completed the 6-month questionnaire, 83% completed the 18-month questionnaire, and 78% completed the 4-year questionnaire (Firestone *et al.*, 2015). Greater levels of attrition were evident in the ELF cohort. Of the 2,197 women recruited between 2008 and 2012, more than one-half were lost to follow-up, leaving a total sample size of 1,042. Participants were considered lost to follow-up if they failed to submit the initial prenatal questionnaire after at least three follow-up reminders since being recruited to the study. Reasons for non-submission included: (1) attrition/loss to follow-up (81%); (2) withdrawal from the study (12%); (3) missing information (1%); and (4) additional reasons (4%), such as moving abroad, miscarriage, nonviable pregnancy, or infant death. Significantly, out of the retained sample (N=1042), more women (55%) opted for a postal survey option, compared to the 45% of participants who chose to complete the study online (Firestone *et al.*, 2015). Although the attrition rate observed in the ELF cohort is alarming, it is imperative to note that this study commenced more than 10 years ago, prior to the advancements in web-based survey administration discussed above. The gradual shift towards online-only methods is ongoing and techniques are ever-evolving (Salvador *et al.*, 2020). The forthcoming GenV study will likely shed further light on the feasibility of online-only data collection because, following initial face-to-face recruitment, GenV plans for most of the contact with families to occur virtually (Nkyekyer *et al.*, 2021).

### 2.2.3 Smartphone apps

Since the Apple iPhone was launched in 2007, the smartphone market has grown exponentially (Ratan *et al.*, 2021). According to recent data, smartphone usage in Ireland is at 95% for those aged 16–29 years, 96% for those aged 30–34 years, and 70% among those aged 45–59 years (Gibney & McCarthy, 2020). With innovations in technology and mobile communications, smartphones are progressively recognised as handheld computers, rather than phones, owing to their strong on-board computing capacity, large screens, and capacious storage space (Spittle *et al.*, 2016). What distinguishes smartphones from previous generations of mobile phones is their capability to run third-party apps provided by companies/developers independent of the smartphone manufacturer. This feature, in conjunction with data connectivity and an array of sensors, makes smartphones an ideal means of collecting data from research participants, especially in their natural environment (Stone & Skinner, 2017).

The sensing features of smartphones can be used to facilitate more sophisticated electronic data capture. For instance, data from sensors can be implemented to automatically trigger the collection of data, and to passively record complex behaviours without necessitating active contribution from the user (Stone & Skinner, 2017). Examples of data that can be passively collected include geolocation and physical movements, online search behaviour and browser history, screen time and app usage, and call and text message logs. Unlike surveys, which depend on self-report, passive data collection has the capacity to record a higher frequency of data, alleviate participant burden, and decrease the risk of measurement error (for example, by eliminating recall errors and social desirability) (Keusch *et al.*, 2019). Yet, while traditional surveys enable participants to control what data are shared with the researcher via a reliance on self-report, passive data collection does not facilitate such curation prior to data being shared (Keusch *et al.*, 2019). Accordingly, the passive collection of traceable, real-world data necessitates an in-depth consideration of ethical and legal compliance, alongside ways to safeguard participant privacy and to certify data security and data protection at all times (Seifert *et al.*, 2018). Giving participants the option to explicitly select which kinds of data they are willing to share in advance (i.e. à la carte consent) could enhance their determination to participate, compared to just providing the option to share all or none of their data. Similarly, participants could be given the opportunity to review, edit, and remove certain data before transmitting them. Both tactics would empower participants to decide to share some, but not all, of their data via passive data collection (Keusch *et al.*, 2019). Passive data collection can also be used in conjunction with traditional survey data to provide a more objective and richer insight into participants' lives (Keusch *et al.*, 2019). A study conducted by Faherty *et al.* (2017) used a novel smartphone app to monitor the daily mood of pregnant women with elevated depression symptoms, who consented to participate in passive data collection. In addition to mothers recording their mood when prompted by a push notification, the smartphone's velocimeter was used to passively collect movement data (total distance travelled by foot) and radius of travel (by car, bike or public transport) per day. The smartphone app analysed this information, combined with metadata concerning the number and length of phone calls and text messages, to identify subtle changes in screen time patterns and movement data that might suggest the need for additional support, and, where relevant, delivered alerts to healthcare providers for personalised support. Through ongoing monitoring of the mothers' internal state and environmental interactions, this app could detect fluctuations in both mood and mobility patterns across the course of the woman's day, week, and entire pregnancy (Faherty *et al.*, 2017).

The commercial market for prenatal mobile health (mHealth) apps is also evolving rapidly, with new parents willingly downloading apps to gain support and information related to their pregnancy and baby through mobile technology (Cawley *et al.*, 2020). The wide availability of mHealth apps for the objective and subjective measurement of health behaviours (e.g. steps walked, sleep patterns, medicine use, mood monitoring, heart rate) could offer a useful means of measuring the health of birth cohort members (Fischer & Kleen, 2021). For example, Radin and colleagues (2018) developed the first pregnancy research app, Healthy Pregnancy Research Program, and embedded this within the already existing (and highly trafficked) WebMD Pregnancy App, which is used for the self-management and education of expectant mothers. Through a combination of survey- and sensor-generated data via the app, a total of 14,045 individual surveys and 107,102 daily measurements of sleep, physical activity and heart rate were successfully collected. On average, expectant mothers remained engaged in the research for 59 days, with 45% of those who reached their due date completing the final outcome survey (Radin *et al.*, 2018).

Subject to General Data Protection Regulation (GDPR) compliance, incorporating a study platform into a trusted mHealth app could be a viable means of promoting visibility and capturing an array of longitudinal, objective, and participant-generated information from a large population of new and expectant mothers (Radin *et al.*, 2018). In a similar vein, a central feature of the new ELC-FS in the UK will be the use of an innovative smartphone app, BabySteps (Cramer Development Incorporated, 2022), to assess and record infant development. At the initial household visit, participants will be asked to record a video of themselves interacting with their child and upload this via the BabySteps app. Following this recording, the primary informant will be encouraged to complete numerous activities via the app, including audio recordings of linguistic interactions; tracking key developmental milestones; completing ecological momentary assessments (EMAs) of parental mood states and infant sleep habits; and additional video recordings of the child's behaviour at home (CLS, 2021b). However, although this proposed data collection would provide rich accounts of the child's developmental trajectory and social environment, online video research is met with complex ethical considerations regarding threats to confidentiality, informed consent, and informational risk (Legewie & Nassauer, 2018).

The merit of smartphone apps to record subjective data is also evident in the MCS, wherein participants (aged 14 years) were given the choice between a paper diary, a web-based diary, or a smartphone app to record their daily activities (Jäckle, Gaia & Benzeval, 2017). The web-based diary could be filled out on netbooks, desktops, and laptops, whereas the app was available via smartphones and tablets on both Android and iOS operating systems. In the pilot study, 75% of participants successfully filled out the diary, with most participants selecting the smartphone app option. An analysis of the quality of the time use data (i.e. quantity and types of activities recorded) indicated that the web diary performed best (only 7% of unusable diaries), followed by the app (23% unusable), and the paper diary option (32% unusable) (Jäckle, Gaia & Benzeval, 2018). Following two rounds of usability testing with young people and the pilot study, the decision was made to include the diary element as a component of the age 14 data collection from January 2015 to April 2016. Based on findings from the pilot study, cohort members were first offered the choice between the online and app modes, whereas the paper diary option was offered only to those who were unable (e.g. did not own the required device) or refused to fill out the diary online or via the app (Ipsos MORI, 2016).

Notwithstanding the promise of smartphone applications, it is critical to also acknowledge potential barriers to effectively delivering smartphone-based research. First, smartphones are relatively expensive, and thus ownership of the most up-to-date model is patterned by socioeconomic status (National Economic & Social Council, 2021). Additionally, app development necessitates specialist technical skills, and often results in extensive testing prior to successful implementation. Consideration must also be given to the types of operating systems that will be supported (e.g. iOS versus Android) and how the app’s compatibility with updates to operating systems will be evaluated and guaranteed. Further concerns pertain to the analysis of potentially large and complex datasets yielded by smartphone devices, alongside ethical considerations surrounding data collection in naturalistic settings (Stone & Skinner, 2017).

### 2.2.4 Wearable activity monitors

Recording physical activity poses methodological challenges for longitudinal research. Most large-scale birth cohorts use self-reported measures to assess physical activity levels, which are inherently susceptible to recall and social desirability biases (Gilbert *et al.*, 2017). In the past decade, wearable devices such as accelerometers and pedometers have been increasingly used as objective measurements of physical activity (Silfee *et al.*, 2018). Similar to commercially available personal fitness trackers (e.g. Fitbit), research-grade activity monitors use motion sensors, but these are specifically calibrated for research purposes, and are often equipped with dedicated software to enable data extraction and visualisation (Stone & Skinner, 2017). For example, valid activity monitor data were successfully collected from 6,497 MCS children when they were aged 7 years (from an accelerometer worn on their right hip) (Rich *et al.*, 2013a), and from 4,221 cohort members when they were aged 14 years (from a wrist-worn accelerometer) (Gilbert *et al.*, 2017). Accelerometers have also been used to objectively measure physical activity during pregnancy; the antenatal wave of the 2015 Pelotas birth cohort study (Brazil) successfully collected accelerometer data from more than 2,000 pregnant women, with 7 days identified as the optimal amount of time required to obtain reliable data from the wrist-worn devices (da Silva *et al.*, 2019, 2018).

Accelerometers can capture a wide variety of movements, alongside identifying different intensity levels of activities. However, it can sometimes be challenging to recognise the exact type of activity the user is engaging in; certain physical activities (e.g. cycling) are more difficult to detect in the case of wrist-worn trackers (Gilbert *et al.*, 2017). Moreover, wearing a visible fitness tracker may result in reactivity, as participants modify their usual exercise habits in direct response to their awareness of being observed (Niemelä *et al.*, 2019). Research-grade accelerometers are also prohibitively expensive and involve considerable training for data collection and analysis (Dominick *et al.*, 2016). In the case of the MCS, as only a limited number of the expensive research-grade accelerometers could be purchased, they were reused throughout the MCS fieldwork period to ensure a broad coverage of all eligible cohort members. Of the total devices that were placed with cohort members (N=9184), 27% of devices were never returned and 9% were returned broken or containing no valid data (Gilbert *et al.*, 2017). Accordingly, a future birth cohort study wishing to gather objective physical activity data should consider the most appropriate type of device to use. In addition to traditional research-grade accelerometers, there are now a wide variety of commercially available personal fitness trackers (e.g. Fitbit), alongside smartphone apps that can be used to detect physical activity and sedentary behaviour. As Fitbits and smartphone apps offer a sleeker and more fashionable alternative to research-grade monitors, they could potentially enhance participant compliance while continuing to collect valid data. Moreover, these devices could decrease the office turnaround time and staff resource requirements associated with the distribution and redistribution of wearable accelerometers, and could potentially expand sample coverage (Gilbert *et al.*, 2017).

Nonetheless, deploying commercially available devices to participants also raises a host of ethical concerns related to the users' privacy, autonomy, and data protection. For instance, the metrics yielded by Fitbits are equivalent to 'personal health data' as defined and regulated by the GDPR (Collins & Marassi, 2021).

### 2.2.5 Biomedical data

Large-scale birth cohorts are increasingly interested in collecting biological samples which, in conjunction with survey data, could enhance understanding of the biological mechanisms of disease (Bailey et al., 2017). Early exposures throughout pregnancy and childhood can permanently alter the body's structure, physiology, and metabolism, subsequently promoting the risk of disease long after the initial exposure has occurred, even across generations. Accordingly, a core objective of many pregnancy and birth cohort studies is to ascertain the influence of early environmental exposures by obtaining biological samples at multiple time points across the child's development (Shamsipour et al., 2020). For instance, Larsen et al. (2013) reviewed European pregnancy and birth cohort studies carried out between 1980 and 2013 (N=56) and identified many cohorts which collected biological samples, including maternal whole blood (43%), maternal serum/plasma (52%), maternal DNA (45%), breast milk (30%), child whole blood (57%), child serum/plasma (60%), child DNA (59%), umbilical cord blood (59%), and paternal DNA (21%). The analysis of these longitudinal biomedical data could serve to promote aetiological knowledge of a variety of childhood diseases (Shamsipour et al., 2020).

The long-term storage of biomedical data is an important component of epidemiological research (Rønningen et al., 2006). Globally, several well-resourced birth cohort studies have established biobank repositories of biological specimens to provide for current and future analyses of the interplay between genetic, lifestyle and environmental influences on children's health and well-being (Townsend et al., 2016). Careful planning and secure arrangements for processing and long-term sample storage are imperative to get the maximum benefit from these sample repositories. Biobank sample quality is contingent on a variety of predetermined considerations, such as accurate sample identification (e.g. use of pre-printed barcodes), safeguards against contamination at sample collection, fast processing and quick transit to storage facilities, long-term storage at optimal temperatures, and storage in a format that facilitates their retrieval for future analyses (Jones, 2009).

Within the Irish context, the Cork BASELINE birth cohort study, established in 2008, was the first ever longitudinal birth cohort study in Ireland to carry out extensive biobanking. Biomedical data were collected from 2,137 families (mothers, infants, and fathers) at 15 and 20 weeks' gestation, at birth, and when the child was aged 24 months and 5 years. Biological samples obtained included maternal blood and DNA, paternal DNA, umbilical cord blood and DNA, and paediatric venous blood sampling at 24 months and 5 years. Following collection, samples were immediately stored at 5 °C; subsequently processed, aliquoted and barcoded within 3 hours of sample collection; and then stored at -80°C for future analysis (O'Donovan et al., 2015). Regarding incidental findings, any children who demonstrated issues regarding the study topics (e.g. eczema, allergies, diabetes, developmental issues) were examined more closely by the study team. Parents were invited to the research centre for further assessment if any concerns regarding their child's health and development arose. For example, children's complete blood count tests were each reviewed by the study's paediatrician, and parents were notified if there were any concerns about the blood test results (BASELINE, n.d.).

Over the course of the 2010s, the advancement of high-throughput biological screening and 'omics' technologies (genomics, epigenomics, metabolomics, lipidomics, and proteomics), as applied to large biobanks, has facilitated scientific research at scale (Bone et al., 2020; Yang et al., 2011). For example, at both an individual and population level, biobanks enable researchers to explore the associations between adverse pregnancy/birth outcomes, environmental exposures (e.g. climate, diet, infection), and genome and phenotypic characteristics (Bone et al., 2020).

Notwithstanding the scientific value of biobanks, research using stored biological samples from children is fraught with ethical considerations. A major ethical difference between paediatric biobanks and adult biobanks is the issue of consent. In the case of birth cohort studies, children become study participants before birth/from very early infancy and do not consent to the research themselves. This poses important questions about the scope of parental consent, the right of the child to assent or dissent to biobank research, and whether re-consenting processes should occur when participants reach a certain age (Hens Lévesque & Dierickx, 2011).

To explore how paediatric biobanks deal with core ethical and legal considerations, Ries, LeGrandeur and Caulfield (2010) carried out semi-structured interviews with investigators from six large-scale birth cohort studies: Born in Bradford (England); the Canadian Healthy Infant Longitudinal Development study (Canada); Copenhagen Prospective Studies on Asthma in Childhood (Denmark); Generation R (Netherlands); Étude Longitudinale Française depuis l'Enfance (ELFE; France); and the National Children's Study (NCS; United States of America). All six studies required informed consent from mothers, and paternal consent was also sought if the father was present and willing to take part in the research. Each of the studies varied in breadth of consent; however, no study requested 'blanket' consent for the use of biological samples for all future research endeavours. Corresponding to ethical guidelines and data protection legislation, each of the studies requested specific, informed consent for each future use of participant data. The ongoing contact with participants throughout the course of the studies provided investigators with opportunities to refresh consent. Ongoing contact also facilitated additional consent for new research activities not anticipated when the initial consent was obtained. The cohorts also had procedures in place to obtain assent from children, sometimes aged as young as 7 years, and additional plans were in place to obtain consent from participants once they reached the legal age of maturity. Moreover, firm data protection measures were implemented and access to any identifiable data was restricted across the six studies. Concerning incidental findings, none of the studies planned to report results of unknown clinical significance to participants, but they varied in their intentions to share clinically meaningful results of more routine tests and measures with the participants. Finally, clear withdrawal options were offered to all participants at any stage of the research. Participants had the option to entirely withdraw all biobank data, or they could opt to withdraw from future contact but consent for the continued use of previously collected specimens (Ries, LeGrandeur & Caulfield, 2010). As these studies predate the GDPR, it should be noted that blanket consent is no longer permissible under the GDPR; consent must be specific and outline the precise purpose of the data processing (GDPR.EU, 2018).

### 2.2.6 Data linkage

Linking birth cohort data to population-based databases (e.g. hospital admissions, birth and death registration data) could provide a richer and more complete insight into the health and well-being of the population at large. Several countries (e.g. the Nordic countries and Australia) have a long-standing practice of incorporating administrative data alongside primary data collection, with such data linkage dating back to the 1980s (Zylbersztejn, Gilbert & Hardelid, 2020). Recently, the increased usage of existing data resources through improved data linkage practices has come to be viewed as one of the most cost-effective ways of supporting longitudinal research (Green *et al.*, 2015). Linking administrative data to survey responses can result in shorter interviews/questionnaires and reduced respondent burden (Al Baghal, 2016; Sakshaug *et al.*, 2012). Furthermore, if participants are lost to follow-up, data linkage (with prior consent) can offer missing information on outcomes such as hospital admissions (Skilton, 2016). Combining multiple datasets also enhances the quality of data by enabling the consistency of data to be checked, and possible missing information to be filled in (Green *et al.*, 2015).

Nevertheless, linking detailed administrative data to equally detailed birth cohort data can be a complicated and lengthy process. Consent for linkage should be obtained from data controllers and from study participants. Data linkage should be conducted securely and precisely, and the resultant linked dataset should ideally be completely anonymised (Tingay *et al.*, 2019). When providing linkage consent in longitudinal studies, participants can often be inconsistent in the consent choices they make, with consent rates varying between the type of data linkage request (e.g. education, health, economic records) and over time (Jäckle *et al.*, 2018). A recent investigation of rates of consent to different types of administrative data linkage among young people partaking in Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC) observed that, although consent rates were above 80% in Waves 6 (aged 14–15 years) and 7 (aged 16–17 years), a slight decline in consent rates was observed between the two waves. Moreover, congruent with previous findings that linkage consent is typically lowest for economic records, LSAC participants' consent rates were lower for income support data compared with medical records (Bandara *et al.*, 2019).

Besides consent, a further challenge in generating reliable linked data is the existence of accurate identifiers that can be used to link the same individual across multiple data sources. If possible, a single unique identifier (e.g. the Personal Public Service Number (PPSN) in Ireland) should be documented across all combined datasets to facilitate a straightforward linkage between sources. Yet, in reality, such an identifier is seldom available, especially when linking data across different sectors (e.g. health, education), which often results in a certain degree of error or missing data (Harron, Doidge & Goldstein, 2020).

Regarding the potential for data linkage in Ireland, the current data infrastructure in the country's health and social care sector is extremely fragmented, with major gaps and isolated silos of information which inhibit the secure, efficient transfer of information (HIQA, 2014). Although multiple policy frameworks have outlined ambitions for the development of health information systems (HIS) via the adoption of vital tools such as individual health identifiers (IHIs) and electronic health records (EHRs), implementing these at a national scale has yet to be achieved. Policy-makers have acknowledged that Ireland seriously lacks a HIS comparable with many of its European peers. For example, Denmark is often recognised as a world leader in HIS and eHealth. Denmark's provision of a centralised identifier that tracks every individual contact with the healthcare system has facilitated world-leading epidemiology research, as its extensive network of medical databases (which record data from birth) provides researchers with extensive data linkage options (Walsh, Mac Domhnaill & Mohan, 2021).

Recent public opinion in Ireland indicates that the majority of citizens would support the development of eHealth and data infrastructure on par with our European counterparts. In a national survey carried out by the Health Information and Quality Authority (HIQA), most people in Ireland (86%) stated that they would be happy for data about their health to be obtained, used and shared electronically via a national patient portal (HIQA, 2021). Based on these findings, and the urgent need for a modern health infrastructure that arose during the COVID-19 pandemic, the feasibility of health data linkage in a future birth cohort study will largely depend on improvements in the current HIS in Ireland, in line with Sláintecare proposals (HIQA, 2021).

Beyond health and social care records, there has been a growing increase in the availability of data at aggregated level in Ireland (e.g. via the CSO's StatBank and data.gov.ie) (Hanafin, 2020). The GUI study is conducted under an arrangement based on Section 11 of the Statistics Act, 1993, which allows the CSO to make arrangements with other public bodies for the collection, compilation, extraction, or dissemination of information for statistical purposes. Accordingly, to fully utilise the information provided by GUI participants, the CSO may link the data from this survey to additional data held about survey participants or the areas in which they live. Operating under the strictest controlled procedures, this linkage would only be carried out for statistical purposes as per the Statistics Act, 1993, and any outcomes of the analysis would not include any identifiable information (CSO, 2022a). As part of the CSO's leadership role of the Irish Statistical System (ISS), policy-relevant research projects (known as 'pathfinder projects') are conducted in partnership with other public sector bodies (CSO, 2022b; Hanafin, 2020). Pathfinder projects entail the integration of secondary datasets hosted by Government Departments and agencies with existing administrative data held by the CSO in order to yield aggregated analysis and outputs. This data linkage for statistical purposes is conducted by the CSO on pseudonymised datasets and is carried out in accordance with the Statistics Act, 1993 (Hanafin, 2020). For example, consent was sought by the GUI study team from members of GUI Cohort '98 to link to their Central Applications Office (CAO) data regarding all applications to higher education, college and university courses in the Republic of Ireland (ISSDA, 2018). Of the 6,216 respondents, 3,203 (52%) gave consent to access their CAO records, and linkage was successfully achieved by the CSO for 3,061 cases (ISSDA, 2018). In relation to a future birth cohort study in Ireland, a pathfinder project could be developed whereby individual-level data from the cohort members are linked with administrative data from a broad range of public sector bodies concerning the individual participants (e.g. the Revenue Commisioners, the Department of Social Protection, the Higher Education Authority) and other issues of relevance (e.g. characteristics of their local area such as deprivation/affluence, population density) (Hanafin, 2020).

### 2.2.7 Conclusion

More than a decade after the initial launch of the GUI study, contemporary data collection innovations present a broad menu of considerations for the development of a new GUI birth cohort. Substantial technological advancements (e.g. web-based surveys, smartphone apps, fitness trackers) facilitate the remote collection of birth cohort data at scale. Yet, switching to a fully digital means of data collection is met with possible limitations, such as lower response rates and an increased risk of measurement error and response bias in the absence of a trained interviewer. Internationally, birth cohort studies have also witnessed a growing interest in the collection and long-term storage of biomedical data to shed light on the aetiology of common childhood diseases and developmental outcomes. For instance, through a combination of longitudinal survey data and genetic information (collected via simple, non-invasive saliva sampling during Wave 6 of the MCS), researchers can explore how genetic and environmental influences shape human development across the life span (DNA from the Wave 6 data is currently being extracted for later genotyping for future research) (Fitzsimons *et al.*, 2021). Although the GUI study is not exclusively epidemiological in focus, the collection of biological samples might be considered in the development of a future birth cohort to enrich survey data across the broader domain of physical health and development. Finally, while the GUI study has successfully utilised administrative data across the sectors of education (e.g. the CAO points system at age 17/18 years of GUI Cohort ’98) and social protection (e.g. use of the Child Benefit register for sampling and inter-wave tracing of GUI Cohort ’08), the impressive integration of healthcare records in several international birth cohort studies accentuates the gaps evident in Ireland’s HIS. Subject to the implementation of proposed improvements in Ireland’s data infrastructure (e.g. IHIs, EHRs), data linkage could be a cost-effective and efficient means of supporting future longitudinal research in Ireland.

## 2.3 Case studies

To complement the foundational study (GUI), this section will examine three additional named birth cohort studies which exhibit innovative data collection techniques: the MCS, the Danish National Birth Cohort (DNBC), and Étude Longitudinale Française depuis l'Enfance (ELFE). Each of the four case studies will be presented under the following subsections: (1) domains of data collection, (2) modes of data collection, and (3) data collection instruments (including data linkage/biomedical data).[16]

### 2.3.1 GUI Cohort '08

#### 2.3.1.1 Domains of data collection

The GUI study was developed to describe and analyse what it means to be a child growing up in contemporary Ireland via an examination of the factors which contribute to and undermine Irish children's well-being (Thornton *et al.*, 2013; Williams *et al.*, 2010). Placing the child at the very centre of this study, the GUI study adopted a multi-informant approach to explore a wide range of child outcomes across three domains: (1) physical and health development, (2) education and cognitive development, and (3) socio-emotional development (including family relationships). To contextualise the outcomes across these broad domains and ascertain how infants' lives differ across varied sociodemographic contexts, the GUI study also collects a vast array of data on the household, parental characteristics (e.g. family income, social class, educational attainment), relationships within the household, and neighbourhood factors (Murray *et al.*, 2020; Williams *et al.*, 2010). Being longitudinal in nature, the GUI study has followed this nationally representative sample across childhood to collect follow-up data when they were aged 3 years (2011), 5 years (2013), 7/8 years (2016) and 9 years (2017–2018), with the most recent wave of data collection completed in June 2022 when the children in the cohort were aged 13 years (GUI, 2022; McNamara *et al.*, 2021). A supplementary online COVID-19 questionnaire was administered in December 2020, and the cohort participants were

---

[16] A comprehensive overview of other international longitudinal birth cohort studies and their specific domains and data topics, including at different waves and different ages of the cohort participants, can be found at http://www.birthcohorts.net. It is important to note that while this website identifies many studies, it is not an exhaustive repository of all longitudinal birth cohort studies worldwide.

recently revisited when they were aged 13 years (using a combination of telephone interviews and web-based surveys). This longitudinal approach has enabled the research team to track developmental trajectories across time; for example, the GUI study's design allows for examining which factors have the strongest impact on children's outcome trajectories from infancy to middle childhood (Williams *et al.*, 2010). To facilitate the discernment of longitudinal patterns, domain topics were repeated at multiple waves in age- and stage-appropriate ways. Supplementary questions were also included across data collection waves to reflect emerging topics. For example, the 2017–2018 wave explored issues particularly relevant to 9-year-olds, such as Internet/computer screen time (Murray *et al.*, 2020). Moreover, a new domain of economic and civic engagement was introduced from age 17/18 (Wave 3, Cohort '98) to reflect the increasing relevance of these topics as the cohort moves into adulthood (McNamara *et al.*, 2020).

### 2.3.1.1.1 Physical health and development

Eight main themes comprised the domain of physical health: pregnancy/prenatal care, child's birth, child's healthcare utilisation, nutrition/diet/breastfeeding, child's physical activity levels/exercise, child's physical development, physical measures, and parental health and lifestyles. Between Waves 1 and 2, there was a clear shift in focus from issues surrounding pregnancy, prenatal care, and labour in the interview conducted when the children were aged 9 months to a greater emphasis on health and healthcare utilisation by the time the children were aged 3 years. Moreover, the theme of diet and nutrition moved from an initial focus on early feeding behaviours and timing of introduction to solid foods (Wave 1) to incorporate other elements of the child's nutritional status (e.g. dietary intake, parental feeding style, weight) (Wave 2). In terms of physical development, data on significant developmental milestones also assumed greater significance by Wave 2, such as information on when the child took their first steps alongside details on gross motor skills (McCrory *et al.*, 2013).

### 2.3.1.1.2 Education and cognitive development

The educational/cognitive development domain encompassed three core themes: childcare arrangements, child's education/home learning environment, and child's cognitive development. Although childcare was a focal point of Wave 1, a shift from parental to non-parental childcare would have occurred for many families by the Wave 2 interviews when children were aged 3 years; attendance at preschool and enrolment in primary-level education would have also commenced for many families at Wave 2. With the 3-year-old cohort (Wave 2), cognitive development was also measured directly using items from the British Ability Scales (BAS). For instance, the BAS Naming Vocabulary and Picture Similarities were presented to assess children's verbal and non-verbal abilities. This addition demonstrates the means through which the GUI study evolved to consider the emerging developmental pathways of the growing child (McCrory *et al.*, 2013).

#### 2.3.1.1.3 Socio-emotional development

The domain of socio-emotional/behavioural development focused on six main themes: child's relationships; child's lifestyle (habits, routines), play and activities; child's socio-emotional development; family context/parenting; marital/partner relationship; and non-resident partner. By Wave 2, the GUI study placed an increasing emphasis on parenting, perceptions of parental self-efficacy, and discipline styles. Moving beyond a focus on topics such as quality of attachment and early communication skills (i.e. understanding of basic commands, efforts to engage in communication using sounds), measurements of children's temperament and emotional and behavioural outcomes were more formally explored in Wave 2 (McCrory *et al.*, 2013).

#### 2.3.1.1.4 Classificatory and background characteristics

To aid the analysis and understanding of child outcomes, the GUI study recorded details on the child's family and other background characteristics. At both 9 months and 3 years of age, a similar set of sociodemographic data was recorded for each child. However, the 27-month period between Wave 1 and Wave 2 witnessed an unprecedented recession in the Irish economy (beginning in 2008). To reflect these emerging social, political, and economic events, the Wave 2 instruments were adapted to measure the effect of the recession on the families participating in the GUI study. For example, a routed question was used to assess the degree to which the recession had affected the household, with four response categories, ranging from 'a significant effect' to 'no effect at all'. Those who indicated that the recession had impacted on their family were then asked further questions about the specific ways in which the recession had affected the household (e.g. being made redundant, being behind with the rent/mortgage, struggling to pay utility bills) (McCrory *et al.*, 2013).

### 2.3.1.2 Modes of data collection

In addition to collecting data from the children directly, their parents/guardians, teachers, school principal and childcare providers participated to varied extents across the five waves of the GUI study (Hanafin, 2020). In the first and second waves of data collection with Cohort '08, effectively all fieldwork took place at the child's home, wherein the primary caregiver (typically the child's mother or maternal figure) and secondary caregiver (typically the spouse or partner of the primary caregiver) were the main informants (Quail *et al.*, 2011b). The main interviews were conducted with each adult via computer-assisted personal interviewing (CAPI), whereas questions pertaining to more sensitive topics were included in a self-completion module on the laptop (CASI).

#### 2.3.1.2.1 Computer-assisted personal interviewing

The main questionnaires were administered by interviewers using a laptop. Interviewers read each question aloud as it appeared on the laptop screen and recorded the participant's response in the space provided. Responses were primarily recorded by entering the relevant number associated with the chosen option. Where applicable, interviewers could record responses by clicking on the computer mouse or by entering free text. Moreover, once each interview questionnaire was completed, it was 'locked down' to prevent the questionnaire being reopened by the interviewer. All questionnaires were programmed using Blaise®. This programme enabled the routing of questions (e.g. skipping non-applicable questions). Blaise® also facilitated the inclusion of hard and soft cross-variable and range checks to notify interviewers of improbable or impossible answers, or of any conflicts between answers. Throughout the course of the interview, participants were presented with an extensive range of prompt cards with the available answer options for each question. These prompt cards were especially relevant for longer lists of response options or items in a scale. Once completed, interviews were outputted from Blaise® as ASCII files, encrypted, and uploaded to a dedicated server in the Economic and Social Research Institute (ESRI) by the interviewers through a telephone line (Thornton *et al.*, 2013). Data were then decrypted and rebuilt to produce a Statistical Package for the Social Sciences (SPSS) file for the purpose of preliminary data analysis. In addition to encryption of the data in transfer, each of the laptops used was secured with 256-bit encryption (McCrory *et al.*, 2013; Thornton *et al.*, 2013).

#### 2.3.1.2.2 Computer-assisted self-interviewing

A CASI format was used for the completion of the sensitive supplements asked of adults, as well as for sensitive questions asked of child and youth participants, where applicable. The GUI team recognised the advantages associated with using self-administered questionnaires (or components of questionnaires) to enquire about sensitive topics and acknowledged the potential benefits of CASI techniques compared to traditional paper-and-pencil surveys. For instance, computerised assessment can enhance the efficiency and accuracy of data collection because automatic branching reduces participant burden and ensures the participant navigates the questionnaire in the intended manner. Range rules and consistency checks help reduce the likelihood of rogue answers (Thornton *et al.*, 2013). Moreover, CASI increases the perception that data remain confidential, reducing participants' embarrassment and enhancing their willingness to disclose sensitive information (Thornton *et al.*, 2013). For the more sensitive questions, the interviewer passed the laptop to the respondent and helped them in completing a series of sample questions designed to familiarise participants with the different types of response formats (e.g. discrete, Likert scale, open-ended, date format). Once comfortable with the demands of the CASI procedure, respondents took control of the laptop and proceeded with the questionnaires on a confidential self-completion basis. Upon completion of the CASI, the questionnaire was shut down and locked, so that neither the interviewer nor respondent could access the sensitive sections of the questionnaire. Respondents were alerted to this via a prompt on the laptop screen. Throughout the course of the CASI procedure, the interviewer was readily available to provide instructions and assistance if required (McCrory *et al.*, 2013; Thornton *et al.*, 2013).

### 2.3.1.3 Data collection instruments

To assess the three primary domains and supporting themes, the GUI study applied a set of data collection instruments appropriate for the age of the children at each wave of data collection.

#### 2.3.1.3.1 Questionnaire structure

For Wave 1 (9 months), the main questionnaire comprised 12 sections, each of which was divided into modules of questions covering sociodemographic characteristics and the three broad domains of physical and health development, education and cognitive development, and socio-emotional development (including family relationships) (see Table 2.1). The primary caregiver CAPI questionnaire consisted of standardised scales to assess psychological, developmental, and behavioural processes, combined with more specific, descriptive questions regarding health, family context and background, parenting experiences, and everyday behaviours (Thornton *et al.*, 2013). A common sensitive supplementary questionnaire was also administered to both the primary and secondary caregiver. The questions were identical for both parents, apart from questions pertaining to fertility and pregnancy, which were asked only of the biological mother. As the questions asked about more sensitive topics, they were contained in a separate module for the respondent to self-complete using a CASI format (Thornton *et al.*, 2013).

Regarding Wave 2 (completed when the children were aged 3 years), the primary caregiver questionnaire was again administered using a CAPI format, and comprised 11 modules to complement and expand upon the data collected at Wave 1 (see Table 2.2). Each module related to sociodemographic variables or mapped onto one of the primary GUI domains. As was the case in Wave 1, a follow-up supplementary sensitive questionnaire was self-completed by both the primary and secondary caregivers in the home as part of the household interview in Wave 2. For a more detailed outline of the precise scales and measures contained in each module, see Appendices 2A.

### 2.3.1.3.2 Standardised scales

As shown in Appendix 2A, a series of standardised scales were administered across Waves 1 and 2 of the GUI study and scored by the study team according to the protocols provided by the scale authors (Thornton *et al.*, 2013).

### 2.3.1.3.3 Anthropometric measurements

At Wave 1, the trained interviewer recorded weight and height measurements of the primary and secondary caregivers, and the weight, length, and head circumference measurements of the 9-month-old infant. To measure height, a Leicester portable height measure was used. With a range of 0.00–2.07 metres, the Leicester measuring stick gives measurements in imperial and metric units; the GUI interviewer recorded height to the closest millimetre. To record adult weight, medically approved seca 761 flat mechanical scales were used. These scales record weight on a metric scale only and have a capacity of 150 kilograms. To record the infant's weight, seca 835 portable electronic scales were used; these scales have a capacity of 50 kilograms. To measure the length of the infant to the nearest millimetre, the seca 210 measuring mat for infants and small children was used. Infant head circumference was recorded to the nearest millimetre using a 70-centimetre (30-inch) disposable paper tape measure (Thornton *et al.*, 2013). In line with standard practice, three separate measurements of head circumference were taken by the interviewer over the course of the household visit (Thornton *et al.*, 2013). At Wave 2, the same instruments were again used to record children's and adults' anthropometric measurements; children's head circumference was not recorded at this wave of data collection (McCrory *et al.*, 2013).

### 2.3.1.3.4 Cognitive ability measure

Whereas the Ages & Stages Questionnaires (ASQ) was administered at Wave 1 to measure developmental milestones (e.g. communication, gross motor, fine motor, problem solving, personal/social development) via parent reporting (Quail *et al.*, 2011b), children's early cognitive abilities were not formally measured until Wave 2. The BAS were selected for use at this stage because they provide a direct measure of cognitive ability among preschool-aged children. Considering the time constraints of the household visit, it was not possible to administer the entire Early Years battery of the BAS; instead, two of the main scales (Naming Vocabulary and Picture Similarities) were chosen to derive a measure of verbal and non-verbal ability among the 3-year-old cohort. The Naming Vocabulary test provides an assessment of children's expressive English language vocabulary, whereas the Picture Similarities test assesses reasoning capacity and problem-solving abilities. For children aged under 3.6 years, these BAS measures produce raw scores that can be transformed to ability scores, percentile ranks, t-scores or age equivalents (McCrory *et al.*, 2013).

**Table 2.1** Summary of instruments used for Wave 1 (2008) of the GUI Cohort '08

| Respondent | Mode | Summary of content |
|---|---|---|
| **Primary caregiver** | CAPI | A. Background information<br>B. Parenting, child's functioning and relationships<br>C. Baby's development<br>D. Baby's habits<br>E. Childcare arrangements<br>F. Siblings and twins<br>G. Prenatal care<br>H. Child's health<br>I. Respondent's health<br>J. Family context<br>K. Sociodemographics<br>L. Neighbourhood and community |
| | Measurements | Height and weight |
| **Primary and secondary caregiver** | Sensitive questionnaire (CASI) | A. Relationship to child<br>B. Current marital status<br>C. Relationship with partner<br>D. Previous relationships<br>E. History of fertility treatment (biological mothers)<br>F. Intention to become pregnant (biological mothers)<br>G. Stress during pregnancy (biological mothers)<br>H. Tobacco and alcohol use during pregnancy<br>I. Parental drug use<br>J. Mental health<br>K. Parental contact with the criminal justice system<br>L. Details about the non-resident parent (if relevant) |
| **Secondary caregiver** | CAPI | A. Stress during pregnancy (biological mothers)<br>B. Tobacco and alcohol use during pregnancy<br>C. Parental drug use<br>D. Mental health<br>E. Parental contact with the criminal justice system<br>F. Details about the non-resident parent (if relevant) |
| | Measurements | Height and weight |
| **Infant (aged 9 months)** | Measurements | Weight, length, head circumference |

Source: Quail *et al.*, (2011b)

**Table 2.2** Summary of instruments used for Wave 2 (2011) of the GUI Cohort '08

| Respondent | Mode | Summary of Content |
|---|---|---|
| **Primary caregiver** | CAPI | A. Household composition<br>B. Child's habits and routines<br>C. Child's physical health and development<br>D. Parental health<br>E. Child's play and activities<br>F. Child's functioning and relationships<br>G. Childcare arrangements<br>H. Parenting and family context<br>I. Sociodemographics<br>J. About you (the primary caregiver)<br>K. Neighbourhood/community |
| | Measurements | Height and weight |
| **Primary and secondary caregiver** | Sensitive questionnaire (CASI) | A. Reasons for people leaving household since Wave 1<br>B. Relationship to child<br>C. Current marital status<br>D. Relationship with partner<br>E. Parental Stress Scale<br>F. Parental self-efficacy<br>G. Currently pregnant (mother only)<br>H. Tobacco and alcohol use<br>I. Parental drug use<br>J. Mental health<br>K. Contact with the criminal justice system<br>L. Details on non-resident parent (if relevant) |
| **Secondary caregiver** | CAPI | A. Introduction<br>B. Parental health<br>C. Parenting and family context<br>D. Sociodemographics<br>E. About you (the secondary caregiver) |
| | Measurements | Height and weight |
| **Child (aged 3 years)** | Measurements | A. Height and weight<br>B. Reasoning (BAS Picture Similarities test)<br>C. Vocabulary (BAS Naming Vocabulary test)<br>D. Gross and fine motor exercises |

Source: Murray *et al*., 2013

### 2.3.1.3.5 Motor development

Wave 2 also witnessed the first direct assessment of the study children's motor development. Informed by a variety of neuro-developmental batteries, six items were used to evaluate children's competencies in the areas of gross and fine motor development via parent reports (McCrory *et al*., 2013). For instance, one item asked about the child's ability to ride a tricycle, and another asked about the child's ability to manipulate toys such as jigsaws and Lego or Duplo bricks. The children themselves were also invited to perform a number of tasks to demonstrate their attainment of developmental milestones in the area of gross and fine motor development. Two tasks, developed to examine gross motor competency, assessed the child's ability to stand on one leg for more than 2 seconds, and their ability to throw a ball in an overhand fashion. Fine motor competencies were examined by instructing the child to draw a straight line once the parent had demonstrated this activity, and by determining whether the child held a pencil in a pincer grip between thumb and forefinger. Interviewers recorded these responses using a simple 'yes/no' answer format (McCrory *et al.*, 2013).

### 2.3.1.4 Data linkage

Longitudinal data linkage is the process of merging longitudinal survey data with a range of other administrative data (Hanafin, 2020).

#### 2.3.1.4.1 Health administrative records

At Wave 1 of the GUI study, biological mothers were asked for consent to access data about the birth of the study child, which is stored within the National Perinatal Reporting System (NPRS). NPRS records, which are stored on all births in Ireland, relate to information on gestational age, birthweight, delivery, and health of the mother and baby at birth (Thornton *et al.*, 2013). Separate permission was also requested to link to data about the child's history on the National Immunization Database (Thornton *et al.*, 2013).

#### 2.3.1.4.2 Personal Public Service Number

The GUI team also sought permission from the Department of Social and Family Affairs (now the Department of Social Protection) to obtain the PPSN (i.e. social security number) of the Child Benefit recipient (typically the mother) from the Child Benefit Register. Primary caregivers consented to the use of their PPSN for: (1) inter-wave tracing, and (2) statistical linkage purposes. Participants had the option to consent for only one of these uses, if preferred. Secondary caregivers were requested to provide their PPSN as well, as this may not be readily available from the Child Benefit Register.

## 2.3.2 Millennium Cohort Study

### 2.3.2.1 Domains of data collection

The MCS is a multidisciplinary study, offering rich insights into developmental trajectories and mobility (e.g. in health, education) over the life course for the UK population overall, and across various subpopulations in the UK (Hanafin, 2020). Seven waves of data collection have occurred thus far: when the cohort children were aged 9 months (2001); 3 years (2004); 5 years (2006); 7 years (2008); 11 years (2012); 14 years (2015); and 17 years (2018). Three online surveys pertaining to the COVID-19 pandemic were also administered to MCS participants from May 2020 to March 2021, and an age 22 years wave (2022–2023) is currently in development.

Somewhat overlapping with the domains of data collection in the GUI study, the MCS presents multiple measures of cohort members' physical, socio-emotional, cognitive, and behavioural development over time, combined with in-depth data on their daily life, behaviour, and experiences (CLS, 2020). A wide range of health-related data have been obtained, in addition to measures regarding child development, cognitive ability and educational attainment. The data also provide a wealth of information concerning the social, economic, and demographic profiles of the cohort members and their families (Connelly & Platt, 2014). Regarding the specific domains covered across Wave 1, data collection encompassed an extensive exploration of household demographics, pregnancy and childbirth, physical health of the parents and child, mental well-being, family income, educational attainment, employment, housing, neighbourhood, parenting, and childcare, alongside lifestyle, attitudinal and relationship factors (Joshi & Fitzsimons, 2016). In addition to follow-up information regarding family composition and sociodemographic characteristics, Wave 2 collected additional information about the child's physical health, behaviour, cognitive development, leisure activities, preschool education, and childcare (CLS, n.d.).

### 2.3.2.2 Modes of data collection

At Waves 1 and 2 of the MCS, the primary modes of data collection were CAPI and CASI, each conducted at the cohort members' homes with the child's parent(s). The full set of CAPI and CASI questions (which usually took 70–75 minutes to complete) was administered to the main informant (typically the mother), and a briefer set of questions (taking about 30 minutes to complete) was administered to the main informant's partner (usually the father) (Joshi & Fitzsimons, 2016). As was the case in the GUI study, the survey instrument was administered by the interviewer via a computer using Blaise® computer-assisted interviewing software. The CAPI and CASI questionnaires for both Wave 1 and Wave 2 of the MCS were divided into modules pertaining to the relevant domain topics (e.g. childcare, health, employment, education). Interviews were carried out with each of the cohort infant's parents (where resident). The mother was generally the 'main' respondent, whereas the father or mother's partner was usually the 'partner' respondent. Several modules of the questionnaire were presented to both the main and partner respondents, while other questions were asked of only one of them (Londra, Calderwood & Millenium Cohort Team, 2006; National Centre for Social Research, 2003).

### 2.3.2.3 Data collection instruments

To facilitate all-island comparisons, much of the questionnaire content in the GUI study was modelled on the instruments previously used with the MCS cohort (Thornton *et al.*, 2013). As depicted in Tables 2.3 and 2.4, there is significant overlap in terms of the domain topics asked of MCS and GUI primary and secondary caregivers across the first and second waves of data collection.

#### 2.3.2.3.1 Scaled measures

The MCS employed a variety of scaled measures to assess children's developmental outcomes, and to gain a greater insight into the physical health and psychological well-being of the primary and secondary caregivers. See Appendix 2B for an outline of the standardised measures used in the first and second waves of the MCS.

#### 2.3.2.3.2 Anthropometric measurements

Unlike the GUI study, physical measurements of cohort members' height and weight did not commence until Wave 2, when the children were aged 3 years.

#### 2.3.2.3.3 Cognitive assessments

Similar to the GUI study, Wave 2 was the first time that cognitive assessments were conducted with the cohort children. The BAS Naming Vocabulary test was administered by interviewers to 3-year-old children during Wave 2 of data collection (Hansen, 2014). The Bracken Basic Concept Scale – Revised (BBCS-R) was also used to examine basic concept development among the cohort children. The BBCS-R assesses the child's understanding of 308 functionally relevant educational concepts via 11 subtests. Only subtests 1–6 were administered during Wave 2. When administered together, these six subtests comprise the Bracken School Readiness Assessment (BSRA), which examines children's recognition of 88 concepts relating to colours (primary colours, basic colour terms), letters (knowledge of upper and lower case letters), numbers and counting (recognition of single- and double-digit numbers, ability to assign a number to a group of objects), sizes (concepts which describe one, two, and three dimensions), comparisons (capacity to match/distinguish objects based on salient characteristics) and shapes (identifying linear shapes and three-dimensional shapes) (CLS, 2020).

**Table 2.3** Summary of instruments used for Wave 1 (2001) of the MCS

| Respondent | Mode | Summary of content |
|---|---|---|
| **Primary caregiver** | **CAPI (main questionnaire)** | A. Non-resident parent<br>B. Pregnancy, labour and delivery<br>C. Baby's health and development<br>D. Childcare<br>E. Grandparents and friends<br>F. Parent's health<br>G. Employment and education<br>H. Housing and local area<br>I. Interests and time with baby |
| | **Sensitive questionnaire, self-completion (CASI)** | A. Baby's temperament and behaviour<br>B. Relationship with partner<br>C. Previous relationships<br>D. Domestic tasks<br>E. Previous pregnancies<br>F. Mental health<br>G. Attitudes to relationships, parenting, work, etc. |
| **Secondary caregiver** | **CAPI (main questionnaire)** | A. Father's involvement with baby<br>B. Pregnancy, labour and delivery<br>C. Grandparents and friends<br>D. Parent's health<br>E. Employment and education<br>F. Interests and time with the baby |
| | **Sensitive questionnaire, self-completion (CASI)** | A. Baby's temperament and behaviour<br>B. Relationship with partner<br>C. Previous partners<br>D. Previous children<br>E. Mental health<br>F. Attitudes to relationships, parenting, work, etc. |

Source: Shepherd *et al.*, 2004

**Table 2.4** Summary of instruments used for Wave 2 (2004) of the MCS

| Respondent | Mode | Summary of content |
|---|---|---|
| **Primary caregiver** | **CAPI (main questionnaire)** | A. Non-resident parent<br>B. Pregnancy, labour and delivery<br>C. Child's health and development<br>D. Childcare<br>E. Grandparents and friends<br>F. Parent's health<br>G. Employment and education<br>H. Housing and local area<br>I. Interests and time with child<br>J. Older siblings |
| | **Sensitive questionnaire, self-completion (CASI)** | A. Child's temperament and behaviour<br>B. Relationship with partner<br>C. Previous relationships<br>D. Domestic tasks<br>E. Previous pregnancies<br>F. Mental health<br>G. Attitudes to relationships, parenting, work, etc |
| **Secondary caregiver** | **CAPI (main questionnaire)** | A. Father's involvement with child<br>B. Pregnancy, labour and delivery (where applicable)<br>C. Grandparents and friends<br>D. Parent's health<br>E. Employment and education<br>F. Interests |
| | **Sensitive questionnaire, self-completion (CASI)** | A. Child's temperament and behaviour<br>B. Relationship with partner<br>C. Previous relationships<br>D. Previous children<br>E. Mental health<br>F. Attitudes to relationships, parenting, work, etc |
| **Interviewer** | **Observations** | A. Home environment<br>B. Neighbourhood<br>C. Parenting behaviour |
| **Child** | **Direct assessments** | A. BAS Naming Vocabulary<br>B. Bracken Basic Concept Scale<br>C. Height and weight<br>D. Oral fluid (saliva) sample |
| **Older sibling** | **Self-completion** | (England only) |

Source: Hansen, 2014

### 2.3.2.4 Data linkage

The MCS has obtained record linkage consent to link to a range of administrative data from relevant Government Departments and agencies across the fields of health and education.

#### 2.3.2.4.1 Health records

At the end of the Wave 1 interview, permission was sought from the primary caregiver to link their survey data to the child's National Health Service (NHS) birth record and maternity episode hospital records. Additional consent was requested to link to the National Health Service Central Register to keep track of the cohort children on this register in the instance of lost contact with the family or the death of a cohort member (Shepherd & Gilbert, 2019). Consent rates for this data linkage in England, Wales, Scotland, and Northern Ireland varied from 83% to 92%, with the highest consent levels seen in Scotland and Wales (Tingay *et al.*, 2019).

At Wave 4 (when children were aged 7 years), permission was sought from 14,043 adults with parental responsibility for their child's routine health records (up until the child's 14th birthday) to be released to the MCS for the purpose of data linkage and subsequent research. These health records were requested from children living in Wales, England, and Scotland to be linked to MCS survey responses within the Secure Anonymised Information Linkage (SAIL) Databank at Swansea University (Bandyopadhyay *et al.*, 2018; Sera *et al.*, 2018; Tingay *et al.*, 2019). The NHS is devolved across the four UK countries (England, Wales, Scotland, and Northern Ireland), with diverse systems for supporting access to routinely collected health data for consented research projects. For instance, provision of data linkage is granted in Scotland by the NHS Information Services Division (ISD), in Wales through the NHS Wales Informatics Service (NWIS) and the SAIL Databank, and in England by NHS Digital. While data controllers in England, Scotland and Wales were approached for the purposes of this study, consent to link to health records in Northern Ireland was not sought due to comparatively small population size compared with the other UK countries (Tingay *et al.*, 2019). Although the NHS Research Ethics Committee only granted permission to link to health data until the child's 14th birthday, the cohort was revisited at 17 years of age to provide verbal informed consent for further linkage to health records and was able to refuse to participate in any element of data collection or to withdraw from the study (Staatz *et al.*, 2021; Tingay *et al.*, 2019). Of the eligible cohort members (N=10757) at age 17 years, 85.7% (N=9214) provided consent for linkage to their NHS records (Fitzsimons *et al.*, 2020).

#### 2.3.2.4.2 Education records

During Wave 4 (age 7 years), the primary caregiver was also requested to consent for data from their child's education records until age 16 years (via the UK Department for Education's National Pupil Database) to be linked to their MCS survey data (Downs *et al.*, 2017). At Wave 7 (age 17 years), further consents to link to education records beyond the age of 16 years were obtained directly from the cohort members themselves (CLS, 2019). Of the 10,757 eligible cohort members at Wave 7, 87.5% (N=9407) consented to education data linkage (Fitzsimons *et al.*, 2020). These linked education data have been used to more precisely explore how children born at the turn of the millennium have fared throughout school, including circumstances related to different academic trajectories and attainment (Hanafin, 2020).

### 2.3.2.5 Biomedical data

In addition to face-to-face interviews and linkage to administrative records, biomedical samples were collected from cohort members at several waves of the MCS. All cohort members participating in Wave 2 of the MCS (N=15382) were invited to provide saliva samples collected within the home setting in order to assess their exposure to common childhood infections (Townsend *et al.*, 2012). A total of 12,473 (81.1% of those invited) saliva samples were returned; 119 were excluded due to incorrect, missing, or non-unique identifiers. An additional 351 samples were collected from twins or triplets; individual participants could not be differentiated because only the household (not the child) identifier had been printed on the barcoded label (Townsend *et al.*, 2012). Of the remaining samples, written consent was obtained for 11,034 cohort members (88.5%). Accordingly, a total of 11,034 samples collected from the 3-year-old children formed the deposit and were tested for total immunoglobulin G (IgG) content and the presence of antibodies to Epstein-Barr virus (EBV), varicella-zoster virus (VZV) and norovirus (Townsend *et al.*, 2012). Secondary analyses using the obtained saliva samples have demonstrated interesting findings, such as the positive association between early life material deprivation and risk of EBV infection among 3-year-olds (Gares *et al.*, 2017), and the relationship between higher family income and higher norovirus-specific IgG titres (Violato *et al.*, 2019).

In 2008, the ‘Every Tooth Tells a Story’ project was carried out as a component of the Wave 4 data collection. When cohort children were aged 7 years, families were invited to participate in this project by posting one or more of their child’s shed milk teeth back to the MCS research team. The primary purpose of the collection was to estimate lead exposure as reflected in the build-up in the children’s teeth, and to compare this with geographical measures of environmental lead (Connelly & Platt, 2014; Parsons & Platt, 2016). However, the response rate for the shed teeth collection was only 21%, with returned samples demonstrating some skew towards more affluent and advantaged cohort families (Parsons & Platt, 2016).

A core feature of Wave 6 (age 14 years) of the MCS was the collection of saliva samples from cohort members and their biological parents during the home visit conducted by trained interviewers. Subsequently, a DNA bank was established from these saliva samples. In total, 23,336 samples are available: 9,259 from cohort members, 8,898 from mothers and 5,179 from fathers. Moreover, there are 4,533 mother, child, and father trios, making the MCS the only nationally representative, population-based study in the UK to contain genetic trios (Fitzsimons *et al.*, 2020). The addition of genetic data to the MCS has the potential to facilitate genome-wide association studies (GWAS) on the basis of study focus traits, trajectories, and familial phenotypes. The MCS DNA bank is anticipated to encourage a variety of investigations into the influence of genetic and environmental factors on human development across the life course, including the use of novel genetic indicators of early life factors and analyses using genetics as a lever for causality (Fitzsimons *et al.*, 2021). The purpose of saliva sample collection was to facilitate the extraction of DNA for later genotyping for future research (Fitzsimons *et al.*, 2020). Thus, at the time of this writing, there are no available examples of practical usage of these genetic data.

## 2.3.3 Danish National Birth Cohort

### 2.3.3.1 Domains of data collection

Focusing primarily on maternal and child health, the DNBC was set up to explore the causal link between early-life exposures and disease in later life, with the aim of revealing opportunities for disease prevention (DNBC, 2020a). The overarching goal of the DNBC was to examine pregnancy complications and diseases in offspring as an outcome of modifiable factors occurring in early life and foetal growth, and their determinants. A particular focus was placed on identifying health issues in childhood and adolescence (e.g. mental illness, asthma, allergies, diabetes, cancer) that may have had potential origin during the prenatal period (Olsen *et al.*, 2001). The domains of health and disease are highly evident across the initial waves of data collection, during which pregnant mothers were asked to provide details on factors that may impact foetal growth and development (e.g. drug, alcohol and tobacco use; environmental exposures; earlier pregnancies and childbirths; fertility treatment; diet; stress; psychosocial stress; exercise; socioeconomic factors; and sleep habits) (DNBC, 2020c).

### 2.3.3.2 Modes of data collection

The children born into the DNBC participate based on their mothers' consent until they are capable of consenting for themselves (no later than at age 18 years). Once invited to the study, all participating mothers were required to return a signed consent form, including their personal identification number (PIN), phone number, and signed permission to link to data contained in medical records and population registers concerning themselves and their unborn child (Olsen *et al.*, 2001). Once informed consent was obtained, three main forms of data were collected for the DNBC: (1) self-reported data on exposures and outcomes; (2) biological data (two maternal blood samples collected during pregnancy and umbilical cord blood from the child at the time of delivery); and (3) data from existing national registers (Olsen *et al.*, 2007). In addition to the administration of two computer-assisted telephone interviews (CATIs) during pregnancy at gestational weeks 12 and 30, two CATIs were conducted with the mother after delivery, when the child was aged 6 and 18 months.[17] The purpose of these telephone interviews was to obtain data on lifestyle factors, minor diseases, use of medicine and environmental exposures related to foetal growth, and lifestyle factors that were not registered in administrative records. On average, interviews were quite brief, lasting approximately 10–15 minutes each (Olsen *et al.*, 2001; Olsen *et al.*, 2014).

---

[17] At recruitment, mothers provided informed consent to participate in the aforementioned data collections. Additional consent was sought from participants upon invitation to any of the additional voluntary sub-studies (e.g. puberty follow-up, COVID-19 study) (DNBC, 2020b).

### 2.3.3.3 Data collection instruments

Unlike the other named birth cohort studies, the initial four CATIs were quite brief, and pertained to the collection of objective data concerning maternal and infant health, environmental exposures, and developmental milestones, alongside self-reported lifestyle factors. As can be seen in Tables 2.5 and 2.6, no standardised measures were administered during the initial waves of primary data collection.

**Table 2.5** Core topics and exemplar questions from the DNBC prenatal questionnaires (gestational weeks 12 and 30)

| Topic | Gestational week 12 | Gestational week 30 |
|---|---|---|
| Earlier pregnancies and childbirths, in vitro fertilisation treatment | "How many times have you been pregnant before?" | N/A |
| Health in general and during pregnancy as well as medical pregnancy examinations | "How would you characterise your health in general? Is it very good, average or not so good?" | "Did you have any internal examination, i.e. the manual inspection of the column of uterus, during pregnancy?" |
| Use of alcohol, drugs, tobacco | "How much did you smoke on average?" | "Did you take other narcotic drugs during pregnancy?" |
| Work situation | "What is your present employment situation? Do you have a job?" | "Are your conditions of employment the same as when we had the first interview?" |
| Home environment | "Housing status? Do you rent, live in a cooperative or do you own your house?" | "Have any rooms in your house been painted during your pregnancy or just before you got pregnant?" |
| Diet | "Have you taken any kinds of vitamins, dietary supplements, fish oils, or herbal medicine while pregnant?" | "Are you a vegetarian?" |
| Sleep | N/A | "Are you sometimes not able to sleep all night due to, for instance, night work, the pregnancy or family problems?" |
| Exercise | "Do you get any kind of exercise during pregnancy?" | "How many times per week do you exercise?" |
| Socioeconomic variables | "How are you being supported financially at the moment?" | "Have you been burdened by your housing situation?" |
| Psychosocial stress | "Have you felt frightened and anxious without reason?" | "Do you worry about your unborn child's health? You can answer: 'not at all', 'a little', 'very much'." |

Source: Andersen, Larsen & Kantsø, 2018a, 2018b

**Table 2.6** Core topics and exemplar questions from the DNBC postnatal questionnaires (6 and 18 months)

| Topic | 6 months | 18 months |
|---|---|---|
| Child diet | "Do you breastfeed your boy/girl now?" | "Within the last month, did he/she have anything else but breast milk substitute in the bottle?" |
| Child development | "Can he/she hold his/her head when you pick him/her up?" | "When was the first time he/she could walk alone without support?" |
| Child vaccinations | N/A | "What was he/she vaccinated against?" |
| Follow-up on prenatal questions | "Which gestational week were you in when you gave birth to your child?" | "Do you have a measurement of cranial circumference taken at another time (after birth)?" |

Source: Andersen & Kantsø, 2018a, 2018b

### 2.3.3.4 Data linkage

#### 2.3.3.4.1 Administrative databases

Denmark has a substantial network of government-maintained, population-based registries, thus providing longitudinal sources of routinely collected administrative, health, and clinical quality data (Schmidt *et al.*, 2019). The provision of a unique PIN to every Danish resident facilitates exact individual-level linkage of all records and lifelong follow-up (Schmidt *et al.*, 2019). The DNBC involves active linkage to such nationwide registers on diseases, demography, and social conditions by means of the unique identification number connected to each of the cohort members. For the data collection to be cost-effective, the DNBC based most health outcome data on routine health registers (Olsen *et al.*, 2001). All women enrolled in the cohort study gave consent for information regarding their pregnancy and birth from the Danish National Patient Register to be linked to the DNBC and gave permission for the establishment of ongoing linkage to the Danish National Patient Register and National Hospital Discharge Register (covering both mother and child inpatients and outpatients from over 98% of Danish hospitals) (Nybo Andersen & Olsen, 2011; Olsen *et al.*, 2001). Furthermore, at the request of the DNBC in January 1997, the National Board of Health (renamed the Danish Health Authority) included measures of head and abdominal circumference and placenta weight as new components of routine birth registration in hospitals (Olsen *et al.*, 2001). The DNBC has additional capacity to obtain data on cerebral palsy, infantile autism, childhood cancers, and diabetes via linkage to specialised disease registers that cover these outcomes. Death and immigration within the DNBC cohort are regularly registered by ongoing linkage to the National Population Registry, whereas clear identification of participants' addresses and phone numbers for follow-up purposes is facilitated through linkage to the Civil Registration System (Olsen *et al.*, 2001).

#### 2.3.3.4.2 Biomedical data

Following oral consent, two blood samples were collected from the mother at her initial (gestation week 8) and mid-pregnancy (gestation week 25) GP visit. At the time of delivery, the midwife collected a blood sample from the child's umbilical cord. The three samples were collected into 10-millilitre ethylenediaminetetraacetic acid (EDTA) vials and sent by post at ambient temperatures to the DNBC biobank, where the blood samples were separated into one plasma aliquot frozen at −150 °C, two plasma aliquots frozen at −30 °C, one buffy coat frozen at −150 °C, and four blood spots on filter paper. These blood samples could be utilised to evaluate biomarkers for dietary intake and nutritional status of the mother and child (Olsen, 2012; Olsen *et al.*, 2007). For instance, the maternal blood samples collected during pregnancy have been previously used to demonstrate moderate but consistent correlations between self-reported dietary intake estimates of selected polyunsaturated fatty acids (PUFAs) and the participants' corresponding plasma biomarker concentrations (Madsen *et al.*, 2019). Accordingly, this study provides validation data for the use of a self-report questionnaire to collect dietary information from DNBC participants.

## 2.3.4 Étude Longitudinale Française depuis l'Enfance

### 2.3.4.1 Domains of data collection

Adopting a multidisciplinary approach, the goal of ELFE was to determine how family, social, economic, and environmental conditions influence children's development, health and socialisation, and to explore the complex interplay between these variables across childhood (Pirus *et al.*, 2010). ELFE investigated developmental trajectories across the broad domains of health, social sciences, and physical environment.

#### 2.3.4.1.1 Health

Regarding health, ELFE places a central focus on children's physical growth (i.e. weight and height) by exploring the associations between growth, sociodemographic factors, environmental exposures, and healthcare consumption (Vandentorren *et al.*, 2009).

#### 2.3.4.1.2 Social sciences

In the field of social sciences, ELFE explores the influence of the family context (e.g. family structure and dynamics), the role of external institutions and interactions (e.g. nurseries, schools, sports clubs) and the impact of wider issues (e.g. poverty, social policy) on the lives and development of the cohort children (Vandentorren *et al.*, 2009).

#### 2.3.4.1.3 Environment

ELFE gathers data concerning the effects of contamination during pregnancy and early childhood by well-known pollutants (lead, mercury, polychlorinated biphenyls (PCBs), pesticides), alongside emerging pollutants (e.g. phthalate), on children's neurocognitive and reproductive development. ELFE also aims to identify the potential short- and long-term risks of exposure to outdoor air pollution for pregnancy outcomes and on children's growth and development of respiratory diseases (Vandentorren *et al.*, 2009). For instance, specific contaminants were measured at the time of childbirth in maternal blood (e.g. PCBs, dioxins, furans, and perfluorinated polybrominated compounds) and in umbilical cord blood (e.g. lead) as part of the biomedical sampling undertaken with this birth cohort (Etzel *et al.*, 2014).

### 2.3.4.2 Modes of data collection

The design of the ELFE birth cohort study consisted of a baseline enrolment interview of mothers at the child's birth (18,329 child participants) wherein retrospective data about prenatal exposures were obtained. This initial interview was followed by a prospective follow-up of the child at multiple time points. This follow-up was based on data retrieval and record linkage from administrative databases (e.g. the National Institute of Statistics and Economic Studies (Institut national de la statistique et des études économiques; INSEE), demographic data, health insurance records, national environment database), in conjunction with multiple waves of cross-sectional surveys (Vandentorren *et al.*, 2009). Cross-sectional follow-up was intensive throughout the infancy and preschool periods (from 2011 to 2015). In addition to regular telephone interviews with both parents when the child was aged 2 months, 1 year, and 2 years, data on infant feeding were obtained every month from ages 2 months to 10 months via web-based or paper interviews (parents were given a choice between the two modes) (Charles *et al.*, 2020). Regarding the mode of data collection utilised during Waves 1 and 2, the telephone interviews conducted when the infants were aged 2 months (2011–2012) and aged 1 year (2012) encompassed the 'main' follow-up surveys. Telephone interviews were conducted by specially trained investigators from INSEE (Vandentorren *et al.*, 2009). When the child was aged 2 months, a trained investigator conducted a telephone interview with the mother (60 minutes) and father (30 minutes); questions pertained to sociodemographic and housing characteristics of the children's families, alongside details about childcare arrangements and the child's diet. When the child was aged 1 year, a standardised questionnaire concerning the child's health and social development was administered via telephone interview to the child's mother (60 minutes) and father (30 minutes) (Hallit *et al.*, 2018; Pirus *et al.*, 2010). The protocol of the telephone interview when the child was aged 2 years was akin to the surveys completed in 2011–2012 when the ELFE children were aged 2 months and 1 year, respectively (INED, 2020b).

### 2.3.4.3 Data collection instruments

See Appendix 2C for a complete overview of the descriptive and scaled measures used during the CATIs conducted when the ELFE children were aged 2 months and 1 year.

### 2.3.4.4 Data linkage

At the outset of ELFE, it was planned to collect data on health and healthcare consumption by requesting access – subject to parental consent – to France's social security data management system (Système National d'Informations Inter-Régimes de l'Assurance Maladie; SNIIRAM) (Charles *et al.*, 2011). However, establishing the procedure whereby the SNIIRAM could pass health records on to the ELFE unit proved to be quite a lengthy legal and administrative process. Setting up data linkage was finalised in 2019 for those families (95%) who had originally consented to link to SNIIRAM data at the maternity units in 2011 (unless parents had since withdrawn consent for data linkage) (Charles *et al.*, 2020; ELFE, n.d.a). Linking the ELFE dataset with health insurance records will be used to gather health information both from active participants and from those lost to follow-up (Charles *et al.*, 2020). This linkage facilitates the retrieval of data related to reimbursed outpatient care (including prescribed drugs) and hospitalisations (including discharge diagnosis) involving the mother during pregnancy and the study child since birth. The raw health insurance data file will only be accessible to the SNIIRAM data manager for ELFE. To guarantee the anonymity of ELFE participants, other researchers will only have access to the pre-processed data (ELFE, n.d.a). In addition to the use of SNIIRAM data, ELFE uses external sources to describe the local environment by linking cohort members' home addresses with data in sociodemographic or environmental surveillance databases (Charles *et al.*, 2011).

### 2.3.4.5 Environmental samples

While mothers were in the maternity unit, 'dust fall collectors' in the form of electrostatic cloths were randomly distributed to consenting mothers at each of the 320 participating maternity units (on the basis of annual number of deliveries) (Rocchi *et al.*, 2015). Mothers were instructed to open the collector and install it in a high position (e.g. on top of the wardrobe) in the child's bedroom when they returned from the hospital. After remaining in place for 2 months, the dust collector was to be enclosed and returned via a prepaid envelope to the ELFE laboratory in charge of microbiological analysis (fungus, mould, etc.) (INED, 2020a). A total of 6,317 dust fall collectors were given to a subsample of mothers, 3,217 of which were successfully analysed (18% of the overall sample at baseline) (Charles *et al.*, 2020; Rocchi *et al.*, 2015)

### 2.3.4.6 Biomedical samples

At inclusion, a subsample of 211 maternity units (chosen based on their expected number of deliveries and proximity to biological treatment and storage units) were invited to undergo biological sampling; 70% (n=154) agreed to take part (Charles *et al.*, 2020). Within the participating maternity hospitals, a total of 4,145 mothers consented to have at least one biomarker measured (approximately 23% of baseline sample) (Dereumeaux *et al.*, 2016). Biological samples were collected in two phases: (1) when the woman was in the delivery room (mother's urine; venous blood sample taken while setting up a blood drip; blood and umbilical cord fragments), and (2) in the days following delivery (the ELFE midwife collected a lock of maternal hair, colostrum from lactating mothers, and samples of the child's meconium and stools taken from nappies). These biomedical samples were obtained, prepared and stored in partnership with the French Blood Agency, which has a network of blood treatment centres across France (INED, 2020a). To give an example of research stemming from this biological sampling, Béranger *et al.* (2018) observed a significant link between concentrations of multiple pesticides in maternal hair samples and children's measurements at birth. Hair samples in particular are relatively simple to collect and store, and pose lower risks of contamination. Hair collected from participating mothers was simply cut, stapled to an index card, placed in an envelope, and transported to be stored at the ELFE biobank, all at ambient temperature (Béranger *et al.*, 2018).

In 2014–2015, when the cohort children were aged 3.5 years, additional non-invasive biological sampling (urine, hair, and stools from children) was performed exclusively from the subset of families from whom biomedical samples were collected at the time of the child's birth. Samples of dust from the home (e.g. from the vacuum cleaner and floor of the room where the child most often played) were also obtained from those who provided samples at inclusion. Specific information was provided to this subgroup of families and informed re-consent was required for both the biological and dust sample collection (INED, 2021). The purpose was to establish a subgroup of ELFE children with available material for assessment of environmental contaminant exposure and health markers at various life stages (Charles *et al.*, 2020). Among the subsample of eligible families, 2,125 children (62% of the 3,415 invited) provided biological samples, and 837 dust samples (81% of the 1,035 invited) were collected from the families' homes (INED, 2021).

## 2.4 Conclusion

This chapter first presented a narrative synthesis of the academic literature on international innovations in data collection practices, specifically the use of technology, the collection and storage of biomedical data, and administrative data linkage. The review was followed by a focus on data collection in the foundation case (GUI) and in three additional case studies: the MCS, the DNBC and ELFE. Relevant to core considerations that might emerge during various stages of the data collection process, the case studies addressed the domains of data collection, modes of data collection, and indicative data collection instruments across domain topics (including data linkage/biomedical data) employed across each of the four case studies. Chapter 3 will focus on design features of birth cohort studies through a review of sampling approaches and recruitment considerations within the international academic literature, and across the named birth cohort studies.

# Appendices

## Appendix 2A

### Standardised scales used in Waves 1 and 2 of the GUI study

**Table 2.7** GUI Wave 1 (aged 9 months, 2008)

| GUI scale | Domain | MCS | ELFE |
|---|---|---|---|
| Condon Maternal Attachment Questionnaire , Quality of Attachment Sub-Scale (Condon & Corkindale, 1998) | Socio-emotional development | ✓ | |
| Infant Characteristics Questionnaire (ICQ) (Bates, Freeland & Lounsbury, 1979) | Socio-emotional development | | |
| Ages & Stages Questionnaires (ASQ) (Squires, Potter & Bricker, 1999) | Physical health and development; education and cognitive development | | |
| Parental Stress Scale (Berry & Jones, 1995) | Socio-emotional development (including family relationships) | | |
| Seven-item Short Form of the Dyadic Adjustment Scale (DAS-7) (Sharpley & Rogers, 1984) | Socio-emotional development (including family relationships | | |
| Centre for Epidemiological Studies Depression Scale (8 items) (Melchior *et al.*, 1993) | Socio-emotional development (including family relationships) | | |

Source: Thornton *et al.* 2013

**Table 2.8** GUI Wave 2 (aged 3 years, 2011)

| GUI scale | Domain | MCS | ELFE |
|---|---|---|---|
| Child-Parent Relationship Scale – Short Form (Pianta, 1992) | Socio-emotional development | ✓ | |
| Parents' Evaluation of Developmental Status (PEDS) (Glascoe, 2006) | Education and cognitive development | | |
| Current Dietary Intake (Sallis, Taylor, Dowda *et al*., 2001) | Physical health and development | | |
| Parental Feeding Style Questionnaire (Wardle *et al*., 2002) | Physical health and development | | |
| 12-Item Short Form Health Survey (SF-12) | Physical health and development | | ✓ |
| Short Temperament Scale for Toddlers (STST) (Prior, Sanson, Smart *et al*., 2000) | Socio-emotional development | | |
| Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) | Socio-emotional development | ✓ | ✓[18] |
| Basic Deprivation Scale (Whelan, Maitre & Nolan, 2007) | Physical health and development | | |
| Four-item form of the Dyadic Adjustment Scale (DAS-4) (c.f. Sabourin, Valois & Lussier, 2005). | Socio-emotional development | | |
| Parental Stress Scale (Berry & Jones, 1995) | Socio-emotional development | | |
| Fast Alcohol Screening Test (FAST) (Hodgson, Alwyn, Hodgson *et al*., 2002) | Physical health and development | | |
| Eight-item Centre for Epidemiologic Studies Depression Scale (CES-D 8) | Physical health and development; socio-emotional development (including family relationships) | | |

Source: McCrory *et al*., 2013

---

[18] Used partially (three questions) when ELFE children were aged 6 months.

# Appendix 2B

## Standardised scales used in Waves 1 and 2 of the MCS

**Table 2.9** MCS Wave 1 (2001)

| Scale | Description | GUI | ELFE |
|---|---|---|---|
| | **A. Assessment of child development** | | |
| Denver Developmental Screening Test (DDST) (Frankenburger, Dobbs & Denver, 1974) | A widely used tool for screening potential developmental delays. Eight items from the **DDST** were used to measure social and communication skills, along with fine and gross motor coordination typical of a 9-month-old infant. | | |
| MacArthur-Bates Communicative Development Inventories (MB-CDIs) | The MB-CDIs are checklists of words and gestures used to assess the development of receptive and productive vocabulary via parent reports. A selection of five items from the UK adaptation of the MB-CDIs were selected to evaluate early communicative gestures. | | ✓ |
| Carey Infant Temperament Scale (Carey & McDevitt, 1977) | To examine the infant's temperament and behaviour, 14 items from the Carey Infant Temperament Scale were used. These items cover four areas: regularity (4 items), approach/withdrawal (3 items), adaptability (2 items) and mood (5 items). | | |
| Infant Control (Avon Longitudinal Study of Parents and Children; ALSPAC) | Derived from the ALSPAC birth cohort study, this measure includes items concerning sleeping, feeding, and crying (including excessive patterns). | | |
| | **B. Assessment of the parent** | | |
| Condon Maternal Attachment Questionnaire (Condon & Corkindale, 1998) | Also used in the GUI study, six items from the Condon Maternal Attachment sub-scale were used to examine mother-to-infant attachment. | ✓ | |
| Parenting beliefs (ALSPAC) | A series of five questions originally developed by the ALSPAC team were used to ascertain parents' attitudes towards parenting (e.g. whether children should be raised in a structured or laissez-faire home environment). | | |
| Golombok Rust Inventory of Marital State (GRIMS) (Rust *et al.*, 1990) | An adapted version of the GRIMS (seven items) was presented to all respondents with a full-time resident spouse or partner to evaluate the overall quality of the couple's relationship. | | |

| Scale | Description | GUI | ELFE |
|---|---|---|---|
| The Malaise Inventory (Rutter *et al.*, 1970) | Parental psychological distress was measured in both parents using an adapted version of the Malaise Inventory (nine items only). | | |
| Rosenberg Self-Esteem Scale (Rosenberg, 1989) | A modified version of the Rosenberg Self-Esteem Scale (six items) was utilised to assess perceived self-worth among both primary and secondary caregivers. | | |
| Life satisfaction | A single item was used to measure general life satisfaction: "On a scale from 0 to 10 how satisfied are you about the way your life has turned out so far?" | | |
| Happiness | A similar measure of happiness was obtained:<br>"Here is a scale from 1–7 where '1' means that you are very unhappy and '7' means that you are very happy. Please enter the number which best describes how happy or unhappy you are with your relationship, all things considered?" | | |
| Locus of control | Control was assessed using a three-item measure of control used in previous waves of UK birth cohort studies (1958 National Child Development Study, 1970 British Cohort Study). | | |
| Social support (ALSPAC) | Designed by the ALSPAC team, social support was assessed using three items concerning emotional, financial and instrumental support. | | |

**Source:** Johnson, Atkinson & Rosenberg, 2015

**Table 2.10** MCS Wave 2 (2004)

| Scale | Description | GUI | ELFE |
|---|---|---|---|
| | **A. Assessment of child development** | | |
| Children's Social Behaviour Questionnaire (CSBQ) | The CSBQ covers three domains: independence and self-regulation (ISR), emotional dysregulation (ED), and cooperation (C). Wave 2 focused on the ISR (five items) and ED (five items) domains. | | |
| SDQ (Goodman, 1997) | The second wave of the MCS used 25 items from the SDQ with both the primary and secondary caregivers. | ✓ | ✓ |

| Scale | Description | GUI | ELFE |
|---|---|---|---|
| | **A. Assessment of parent** | | |
| Life satisfaction | A single item was used to measure general life satisfaction: “On a scale from 0 to 10 how satisfied are you about the way your life has turned out so far?” | | |
| Happiness | A similar measure of happiness was obtained:<br>“Here is a scale from 1–7 where ‘1’ means that you are very unhappy and ‘7’ means that you are very happy. Please enter the number which best describes how happy or unhappy you are with your relationship, all things considered?” | | |
| Locus of control | Control was assessed using a three-item measure of control used in previous waves of UK birth cohort studies(1958 National Child Development Study, 1970 British Cohort Study). | | |
| Social support (ALSPAC) | Designed by the ALSPAC team, social support was assessed using three items concerning emotional, financial and instrumental support. | | |
| Effective Provision of Pre-School Education (EPPE) Project | The MCS used items from the routine/self-regulation scale used in the EPPE Project (e.g. how often do parents read to the child, how much sleep does the child get, how many hours does the child spend watching TV, etc.). | | |
| Straus's Conflict Tactics Scale (Straus & Hamby, 1997) | This scale was presented to the primary caregiver to explore discipline practices and measure possible physical and psychological maltreatment of children. | | |
| Home Observation Measurement of the Environment – Short Form (HOME-SF) (Caldwell & Bradley, 1984) | The HOME-SF is an observation instrument that was used by the MCS interviewer to assess the child’s home environment across a number of subscales (physical environment; emotional and verbal responsiveness of the mother; organisation of physical and temporal environment; and modelling). | | ✓ |
| Child-Parent Relationship Scale – Short Form (Pianta, 1992) | This 15-item scale was used to explore the mother’s feelings and beliefs about her relationship with her child, and about the child’s behaviour towards their mother. | ✓ | |
| Kessler Psychological Distress Scale (K6) (Kessler *et al.*, 2003) | The Kessler (K6) Scale was used to assess levels of psychological distress and depression- and anxiety-related symptomatology among parents. | | |

Source: Johnson, Atkinson & Rosenberg, 2015

# Appendix 2C

## Overview of interview content in the ELFE study (Waves 1 and 2)

**Table 2.11** ELFE 2-month CATI (2011–2012)

| Scale/question | Description |
|---|---|
| List and civil status of inhabitants of residence | The respondent was asked to list and provide demographic details for all persons living in the household. This section was derived from surveys on resources and living conditions (The Statistical Survey on Resources and Living Conditions; SRCV). |
| Child in care (if relevant) | Questions asked of mothers if the study child was not currently living with them (e.g. in foster care). |
| Family situation | Drawn from surveys on resources and living conditions (SRCV), this section asked about marital status and mother's/partner's relationship to the study child. |
| Siblings living outside household | This section asked about the presence of other biological children living outside the home. Where relevant, demographic details for each sibling were recorded. |
| Extended family | This section asked about the study child's grandparents (whether they are alive, resident in the household, etc.) and asked about the level of childcare support provided by the grandparents. |
| Education | Drawn from a previous labour force survey (The Employment Survey, on employment, unemployment and inactivity; EEC), this section asked about the highest level of educational attainment. |
| Employment | Also derived from the labour force survey (EEC), this section asked about the current work situation and employment history of each household member. |
| Housing | Adapted from the INSEE Housing Survey, this section asked about the type of dwelling the family lived in, number of rooms, pollutants and issues related to the local area. |
| Household income | Adapted from a previous survey on resources and living conditions (SRCV), this section asked about sources of income among household members. |
| Living conditions | Also drawn from previous surveys on resources and living conditions (SRCV), this section asked about rent, loans, financial situation, living standard and cost of living, and disposable income. |
| Exposure and environment/pregnancy history | Environmental exposures and mother's behaviour during the neonatal period was examined using the Neonatal Behavioral Assessment Scale (NBAS), the HOME inventory (Caldwell & Bradley, 1984), the French-language version of Q-Sort, and questions developed specifically for the purposes of the ELFE survey. |

| Scale/question | Description |
|---|---|
| Maternity leave and intention to return to work | Questions asked about length of maternity leave taken and date of return to work/intentions to recommence employment. |
| Parent's union status | Asked about the couple's relationship history. |
| Child's health | Respiratory diseases: adaptation of the International Study of Asthma and Allergies of Childhood (ISAAC) questionnaires and, for younger children, the Paris cohort questionnaires.<br>Accidents: adaptation of the French Institute for Public Health Surveillance (Institut de veille sanitaire; InVS) questionnaires on domestic accidents.<br>Asked for self-report details about the infant's weight, height, and head circumference at the time of birth and as per their most recent doctor's appointment. |
| Being a parent | Languages spoken to the child, religion, feelings about becoming a parent/parenthood. |
| Organisation of daily activities | Used the French version of the Generations and Gender Survey (Study of Family and Intergenerational Relations (ERFI); French Institute for Demographic Studies (L'Institut national d'études démographiques; INED)). |
| Relationship between parents | Adapted version of the National survey on Violence Against Women in France (Paris University). |
| Child's diet since birth | The dietary module was designed specifically for the purposes of ELFE by modifying and complementing the questionnaires used in the EDEN and Opaline cohorts. |
| Mother's health and family health history | Asked about the mother's history of medical conditions, illnesses, allergies, hospitalisations, etc., and for details about alcohol, tobacco and/or drug use. |
| Edinburgh Postnatal Depression Scale (Cox JL, 1987) | Used to identify the presence of symptoms of depression and anxiety that are common in women during pregnancy and in the year following the birth of a child. |
| Household caregiving and responsibilities | Asked about division of childcare and household responsibilities. |

***Note:*** The study child's mother (60 minutes) and father (30 minutes) were both interviewed about the same topics as outlined above.

**Source:** INED, 2020b

**Table 2.12** ELFE 1-year CATI (2012)

| Scale/question | Description |
|---|---|
| List and civil status of inhabitants of residence | The respondent was again asked to list and provide demographic details for all persons living in the household. Any changes from Wave 1 were recorded. |
| Child in care (if relevant) | Questions asked of mothers if the study child was not currently living with them (e.g. in foster care). |
| Family situation | Derived from surveys on resources and living conditions (SRCV, INSEE). Where relevant, asked about situation of non-cohabitation or non-permanent cohabitation of parents. |
| Separation and relationship with other parent | If parents were separated, asked about continued contact with partner if he/she does not live in the household. Questions taken from the labour force survey (EEC) and adapted for the purposes of ELFE. |
| Main situation regarding work | Drawn from a variety of previous INSEE surveys, this section asked about the current work situation and employment history of each household member. |
| Education | Drawn from the labour force survey (EEC), this section asked about education history and achievements. |
| Housing | Administered the housing survey (Enquête Logement) from INSEE. |
| Household income | Adapted from surveys on resources and living conditions (SRCV), asked about sources of income among household members. |
| Extended family | This section asked about the study child's grandparents and their level of involvement in the child's life/care at this stage of development. |
| Type of care | This section asked about care arrangements currently in place for the child (e.g. creche, maternal grandparents, minded in the home). |
| French version of the Child Development Inventory (Inventaire du Développement de l'Enfant; IDE) (Duyme *et al.*, 2010) | Administered to explore the child's development of social, self-help, motor, language, and letter and number skills. |
| MB-CDIs (Kern *et al.*, 2010) | Used 11 words from the MB-CDIs to evaluate early communicative gestures. |
| Child's diet | The complementary feeding module was designed for the ELFE survey by modifying and expanding on the questionnaires of the EDEN and Opaline cohorts. |

| Scale/question | Description |
|---|---|
| Child's health | Respiratory diseases: adaptation of the ISAAC questionnaires and, for younger children, the Paris cohort questionnaires.<br>Accidents: adaptation of the French Institute for Public Health Surveillance's (INVS) questionnaires on domestic accidents.<br>Asked about the general health of the child, whether they had received all vaccinations, the presence of any chronic conditions, and hospitalisations and visits to specialists. |
| SF-12 Health Survey (Gandek *et al.*, 1998) | Used to enquire about mother's/father's self-perceived health. |
| Minimum European Health Module (Eurostat-Cox *et al.*, 2009) | A measure of self-reported health, used to enquire about the mother's/father's health and family health history. |
| Information on mother's/father's childhood | Asked about the presence of any speech, learning or behavioural difficulties during the mother's/father's childhood. |
| Educational practices | Asked about how often the parent(s) engage in educational activities/games with the study child (e.g. drawing, reading, singing songs, playing with toys). |
| Values and affiliations | Specifically designed for ELFE, questions pertaining to the parent's involvement in various organisations and the parent's level of trust in different institutions/government organisations. |
| Neighbourhood/local area | Asked about the local area, the services available there, and ways in which the family gets around (e.g. public transport, by car). |
| Cultural goods in the household | Asked about the number of books and CDs, and the technological equipment in the household, including televisions, computers, etc. |

Source: INED, 2020c

# **Chapter 3:** Design features of birth cohort studies

## 3.1 Chapter overview

This chapter will begin with a narrative synthesis of the academic literature concerning the key design features involved in the development of a longitudinal birth cohort study. The importance of a thorough sampling strategy will be discussed, including the need to clearly define the target population, select a suitable sampling frame, and determine the required sample size of participants from the a priori-defined population. Ensuring ongoing representativeness of the birth cohort will also be considered through a review of international best practices in the selection, recruitment, and retention of members from traditionally hard-to-reach groups. An evaluation of the terminated National Children's Study (NCS) in the United States of America (USA) and the Life Study in the UK will then highlight core lessons for recruitment to inform the potential new birth cohort in Ireland. Legal responsibilities to uphold children's data protection rights and ensure child welfare at all stages of the research process will also be discussed. Finally, the review will briefly evaluate four potential sampling frames and their implications for a new birth cohort study within the Irish context. Following this narrative synthesis, the case studies (Section 3.3) will explore how each of the case studies tackled the main design considerations associated with developing a birth cohort. The following topics will be addressed for each case study: cohort age (prenatal versus postnatal recruitment), sampling frame selection, sample size calculation, ensuring representativeness (e.g. reweighting, oversampling, boosted samples, geographical coverage), the timing of additional waves, and retention strategies.

## 3.2 Narrative synthesis of the academic literature

Unlike many research designs, the use of repeated measurements in longitudinal birth cohorts can help ascertain the temporal order of life course events and the direction of association between exposures and outcomes, and, under certain assumptions, can facilitate causal inference (Teague *et al.*, 2018). Nonetheless, developing a birth cohort study is a long-term investment which necessitates careful planning, a high-quality design, and the use of robust data collection methods (Al-

Rifai *et al.*, 2020). Given the lengthy duration and substantial effort involved, the design and maintenance of a potential new birth cohort study in Ireland would require a clear study concept, substantial funding, and precise management at all stages of the development process (Piler, Kandrnal & Bláha, 2017). The aim of this section is to provide a synopsis of the different design features and core considerations involved in the development of a birth cohort study, including the clear definition and selection of the study population, sampling strategies, and recruitment considerations (e.g. hard-to-reach groups, data protection). This section will also highlight lessons for recruitment that have evolved from the cancellation of two large-scale birth cohorts in the UK and the USA. Finally, the review will briefly consider potential sampling frames and their implications for a new birth cohort study in Ireland.

### 3.2.1 Sampling design

When designing a new birth cohort study, the first step should involve identifying a suitable sampling design that will produce a representative sample from which generalisations can be drawn about members of the predefined target population (Canova & Cantarutti, 2020; Sullivan, Joshi & Williams, 2020). Specifying the sampling design in advance is critical; the absence of a rigorous sampling procedure runs the risk of producing samples that do not represent the population at large (e.g. due to selection bias) (Martínez-Mesa *et al.*, 2016).

### 3.2.2 Target sample

Above all, the target population must be clearly identified, relevant characteristics of participants specified, and the eligibility criteria defined (Casteel & Bridier, 2021). Depending on the birth cohort's objectives and the specific hypotheses to be tested, the sample may be selected from the general population or drawn from subgroups of interest within the population (Araujo *et al.*, 2018). The exact geographical coverage of the study should also be determined; whether it will focus on a particular region or state in a country (e.g. GenV, Born in Bradford, Generation R), or whether it will be nationally representative (e.g. GUI, MCS, DNBC, ELFE).

Another important issue to consider when selecting the study population is whether the cohort should be recruited during pregnancy or after the infant is born. The previous GUI Infant Cohort first surveyed participants postnatally when they were aged 9 months. Internationally, birth cohorts primarily engage in two approaches to cohort recruitment: postnatal sampling and/or sampling of pregnant women during the prenatal period. For instance, in a recent review of 111 European birth cohort studies, 66 began enrolment during pregnancy (2 of which began pre-pregnancy), while 45 recruited cohort members at birth or several months afterwards (Pansieri *et al.*, 2020). From an epidemiological perspective, pregnancy is an optimal time to sample participants, as this is a period highly sensitive to maternal influences (e.g. nutrition, exercise, smoking, alcohol use) that can have lifelong health and well-being impacts on the unborn infant (Townsend *et al.*, 2019). In fact, the causal underpinnings of many prevalent diseases in adulthood (e.g. obesity, psychopathology, cardiovascular disease) are believed to have origins in utero and during the early postnatal period (McDonald *et al.*, 2013). Although indirect assessments of prenatal exposures and environmental factors remain feasible after the infant is born, data collection during pregnancy is less susceptible to recall bias and can facilitate the incorporation of prospective repeated objective measures (e.g. physiological assessments, biomedical sampling), in conjunction with maternal self-reports (Goldstein *et al.*, 2021).

Nevertheless, there are distinctive challenges associated with the recruitment and longitudinal follow-up of mother–infant dyads, including the busy schedules of expectant mothers, an unwillingness to undergo intrusive biological sampling, the reluctance mothers may have to participate when they are uncertain of what the outcome of pregnancy will be, and the lack of national listings of pregnant women available in many countries (Goldstein *et al.*, 2021; Manca *et al.*, 2013; Sullivan, Joshi & Williams, 2020). Regarding the feasibility of sampling a pregnancy cohort within the Irish context, the Maternity and Infant Care Scheme delivers an agreed course of care to all expectant mothers living in the Republic of Ireland and is provided by the woman’s family doctor (GP) and a hospital obstetrician (HSE, 2021). Given the GDPR, unless permission was granted to access secure Health Service Executive (HSE) records, the absence of a national database would warrant extensive collaboration with antenatal service providers all over the country. Although pregnancy samples have previously succeeded in certain countries due to strong support from healthcare staff and extensive collection of register data (e.g. DNBC), recent attempts to sample pregnant women have also been met with poor response rates (e.g. the cancelled NCS in the USA and UK Life Study) (Sullivan, Joshi & Williams, 2020).

Therefore, given that the GUI study is not overly epidemiologically focused, alongside the inherent challenges associated with recruiting a pregnancy cohort, recruiting infants postnatally might be the best option in the development of a new Irish birth cohort study. In the original 2001 GUI study design brief, it was acknowledged that recruitment at birth or during pregnancy each have varied implications for the type of data collected, the timing of waves, the availability of sampling frames, and cost. While there are certain measurements that can only be accurately obtained during pregnancy (e.g. ultrasound measurements of the child’s development in utero), beginning data collection as early as possible using a classic birth cohort design can facilitate a child-centred approach to data collection wherein invaluable information about important life events and issues of concern to children can be reliably obtained (Greene *et al.*, 2001). With regard to a future birth cohort, recruiting postnatally would certainly circumvent the potential barriers associated with sampling pregnant women at a national level within the Irish context. Moreover, this would also enable comparisons to be made between GUI Cohort ’08 participants and the next generation of infants growing up in contemporary Ireland. Comparisons between a new birth cohort and the previous Cohort ’08 participants would enable researchers to gather invaluable data about changes in Irish society and how the new generation of Irish children and parents are faring in comparison with their predecessors. Whereas Cohort ‘08 took place during an unprecedented recession in the Irish economy, a new GUI birth cohort study could evaluate, among other unexpected events, the impact of the COVID-19 pandemic on infants and their caregivers, who may have been impacted by the economic and health consequences of the pandemic.

### 3.2.3 Sampling frame

A sampling frame refers to a list of individuals from the target population that can be contacted and potentially recruited into the cohort (Allen, 2017). For example, in a study exploring the quality of childhood cancer care in Denmark, the Danish Childhood Cancer Registry was identified as an appropriate sampling frame because this database contained information on all Danish children aged under 15 years with a diagnosis of cancer (Schrøder *et al.*, 2016). Essentially, “the ideal frame for any statistical survey is an up-to-date and fully comprehensive listing of all elements of the relevant population in question” (Quail *et al.*, 2011a, p. 6). High-quality sampling frames are those that are kept fully up to date and include only one record for each member of the target population; there should be no omissions and no duplication (Mohadjer *et al.*, 2016; Quail *et al.*, 2011a; Watson & Lynn, 2021). Sufficient information (e.g. contact information, sociodemographic variables) must also be contained within the chosen sampling frame to facilitate stratification, recruitment, data collection, weighting, and non-response bias analyses (Mohadjer *et al.*, 2016). For record linkage and follow-up purposes, the ideal sampling frame should include a common identifier for each member of the target population (e.g. PPSN) for the performance of exact linkage and interwave tracing (Burg *et al.*, 2019). Furthermore, the accessibility of the sampling frame must be considered, including both researcher accessibility (e.g. centralised data, availability of the sampling frame in a usable information technology (IT) format) and accessibility of contact information (e.g. addresses, telephone numbers) for recruiting potential participants (Burg *et al.*, 2019).

In the case of large-scale birth cohorts, it must be determined whether the chosen sampling frame aligns with the study’s objectives, and if there are tactics to overcome any identified sampling frame limitations (Martínez-Mesa *et al.*, 2016). For instance, statistically reweighting the survey data can help compensate for any imbalances in the finalised study sample, as compared with the target population, that might have arisen from the choice of sampling frame (Quail *et al.*, 2011a). As will later be discussed in greater detail (see Section 3.3.3.6), the recruited sample of the ELFE birth cohort was drawn from a sampling frame with unequal probabilities of selection (random sample of 349 maternity wards in metropolitan France); thus, weighting methods were employed to adjust for non-representativeness of the ELFE cohort compared to the general French population (Charles *et al.*, 2020).

### 3.2.4 Sample size calculation

Determining the required sample size is another key consideration in the development of a new birth cohort study. Sample size calculations are often directly influenced by the research questions addressed by the study. Yet, it is not unusual for the reverse to arise, whereby the sample size is initially influenced by practical considerations (e.g. time frame, available resources), with feasible research questions being posed thereafter (Golding & Steer, 2009). Although large sample sizes are often attractive, these might need to be traded off against time and/or budgetary constraints (Sullivan, Joshi & Williams, 2020).

The selection of both the exposure and outcome measures of interest will also shape the size of the required sample (and vice versa). Oftentimes birth cohorts that focus on dichotomous outcomes (i.e. whether the cohort member either has or does not have a specified condition) typically possess far greater sample sizes than those with a focus on outcomes that can be assessed on a continuous scale (e.g. physical growth, personality, behaviour, temperament). For example, the DNBC and the Norwegian Mother and Child Cohort (MoBa) study each have a sample size of around 100,000 participants because they are investigating relatively rare exposures (e.g. environmental contaminants, infections) and outcomes (i.e. rare childhood diseases). To achieve sufficient power for investigating rare aetiologies, the clear benefit of large-scale birth cohorts is apparent in the smaller odds ratios that are expected to be statistically significant, even when the actual exposures and outcomes are rare (Golding & Steer, 2009). Conversely, birth cohort studies designed to investigate more common outcomes (e.g. obesity), with a predominant focus on continuous trait measures, may not require as large-scale a sample as those seen in the Nordic studies (Golding & Steer, 2009). For instance, while ALSPAC is an epidemiological birth cohort study, the focus is on relatively common conditions (e.g. eczema, food allergies, asthma) with minimum prevalence of 2% in the population, and thus ALSPAC was deemed sufficiently powered with a sample size of around 14,000 (Golding & Steer, 2009).

Although birth cohorts that are more multidisciplinary in scope (e.g. MCS, GUI) do not necessitate a sample size of 100,000, recruiting a sufficiently large sample remains advantageous in offering the statistical power needed to make robust, policy-relevant inferences (Western *et al.*, 2014). For instance, a large sample size is particularly beneficial when there is a desire to make valid comparisons between specific population subgroups (e.g. migrants, ethnic minorities) or on smaller administrative geographies (Greene *et al.*, 2001). Additionally, the occurrence of significant and systematic attrition over time can severely hinder the generalisability of study outcomes and reduce the statistical power required to detect meaningful effects (Teague *et al.*, 2018). Bigger samples help reduce the impact of attrition by ensuring that even when participants are lost to follow-up, enough participants remain to sustain the research value of the study (Western *et al.*, 2014). Therefore, baseline sample size calculations for effects should acknowledge that the sample size will not remain constant across waves and should take into account the inevitable occurrence of non-response. In this context, Golding and Steer (2009) advise longitudinal researchers to presume that approximately one-half of the sample will actually be available for analytical purposes, and to apply this premise in calculating a conservative estimate of the actual sample size required. As a rule of thumb, it is also suggested that, in most cases, the optimum birth cohort should be designed to recruit the greatest number of participants for which in-depth and precise data collection would be feasible (Golding & Steer, 2009). For example, after accounting for key factors of cost, feasibility, data quality (i.e. minimising bias and ensuring high reliability) and sample attrition over time, the previous GUI study design brief recommended that a high-quality probability sample size for Cohort '08 needed to achieve a sample of at least 10,000 children, or one-fifth of annual births (Greene *et al.*, 2001).

### 3.2.5 Ensuring representativeness

Notwithstanding sample size calculations and the selection of a relevant sampling frame, the actual representativeness of the birth cohort can impact on its capacity to answer proposed research questions, and for its findings to have direct applications for public policy and service improvement (Woolfenden *et al.*, 2016). Differential study participation, such as greater non-participation among socially disadvantaged or ethnic minority groups, could result in an underestimation of the actual prevalence of significant outcomes among these high-risk groups, and restrict applicability of the birth cohort's findings (Woolfenden *et al.*, 2016). However, it is imperative to note that representativeness is not a necessary condition for all types of birth cohort studies; the benefit of a representative sample depends on the main study question and context. Whereas scientific inference of causal relationships in epidemiology (e.g. a disease with a risk ratio under 5%) does not always necessitate representativeness of the general population to be valid, ensuring ongoing representativeness is crucial for policy-oriented cohort studies like GUI (Richiardi, Pizzi & Pearce, 2013).

To ensure representativeness at the sampling stage, recent international birth cohorts have implemented two primary strategies. First, very large samples have been utilised to ensure that a substantial number of children from policy-relevant social subgroups are included (e.g. the DNBC and MoBa each enrolled 100,000 children). This is an effective but costly strategy. In fact, the proposed NCS in the USA aimed to follow a representative sample of 100,000 children but was stalled due to execution issues (e.g. challenges relating to study design and management; rising costs) associated with the magnitude of the sample (Western *et al.*, 2014). The second and less expensive strategy involves oversampling minority groups of interest, and using moderately large sample sizes (e.g. the MCS included almost 20,000 infants but oversampled those from deprived backgrounds and areas of high ethnic minority concentration). Oversampling is beneficial in certifying that sufficient numbers of respondents with characteristics of policy relevance are included in the cohort. Oversampling reduces the impact of the often higher attrition rates among hard-to-reach groups. Yet, for oversampling strategies to be effectively implemented, existing knowledge about the relevant subpopulations and their geographical distribution is essential (Western *et al.*, 2014). To briefly contextualise this within the Irish context, the Traveller community is an extremely marginalised, but traditionally underrepresented indigenous group. Traveller culture places a strong emphasis on living in close proximity to extended family, and so the population density of Travellers tends to be higher in certain cities (e.g. Cork, Galway) and towns (e.g. Tuam, Dundalk, Longford) throughout Ireland (CSO, 2016a). Thus, oversampling children from the Traveller community would likely involve stratification by geographical area and a more targeted strategy.

### 3.2.6 Ensuring participation

With sampling issues considered, recruitment strategies should also be adapted to ensure suitable representation of hard-to-reach and traditionally underrepresented populations. An awareness of cultural norms, practices, and languages is needed in order to foster trust and encourage working relationships with diverse groups. Cultural competence among research staff and field staff is therefore critical for the recruitment and retention of vulnerable populations, such as ethnic minorities, immigrants, and refugee groups (Goldstein *et al.*, 2021). Further recruitment recommendations include the use of personal contact at study enrolment, certifying that those involved in participant recruitment are competent and engaged in the study, engaging in clear and confidential communication with potential participants, and distributing promotional material that is positively valenced (Leung *et al.*, 2013). Furthermore, making study materials (e.g. information sheets, consent forms, questionnaires) accessible and user-friendly, and available in both dominant and non-dominant languages, can promote inclusivity, but can also be met with an increase in cost and resources required (Goldstein *et al.*, 2021).

### 3.2.7 Lessons for recruitment

An extensive follow-up period is imperative in order to determine the multitude of factors that influence children’s health and well-being and also to uncover potential corrective interventions. Yet, the cost associated with designing and operating birth cohort studies over long periods of time is a considerable barrier to their successful implementation (Pansieri *et al.*, 2020). Together with budgetary constraints, a persistent trend, evident internationally, is the steep decline in response to population-based studies (Goldstein *et al.*, 2017). For example, two high-profile birth cohort studies, the NCS in the USA and the UK Life Study, were cancelled despite their substantial expenditure of £800 million and £9 million, respectively (McCarthy, 2014; Nkyekyer *et al.*, 2021; Pearson, 2015). The termination of these two studies is largely attributed to initial challenges associated with participant recruitment and engagement (Nkyekyer *et al.*, 2021). The cancellation of the NCS and the Life Study, each of which aimed to recruit around 80,000 to 100,000 infants, could deter future researchers from developing large birth cohort studies in the coming years (van

Gelder *et al.*, 2020). However, instead of becoming discouraged, valuable lessons for recruitment that have arisen from these stalled birth cohort studies should inform the potential new birth cohort study in Ireland.

### 3.2.7.1 The National Children's Study

The NCS was an ambitious effort to recruit and follow a nationally representative sample of 100,000 children in the USA from before birth until age 21 years (NCS, 2016). In line with the congressional mandate, the overarching goal of the NCS was to elucidate complex interactions between environmental exposures (chemical, biological, physical, and psychosocial) and health outcomes from the prenatal period into early adulthood. To achieve this, the NCS intended to collect and bank serial environmental samples and biological specimens for every mother–child dyad in order to facilitate future assessment of exposures alongside genetic and epigenetic analyses (Hudak *et al.*, 2016). The NCS comprised two distinct but related studies: the NCS Vanguard (Pilot) Study, and the NCS Main Study. The Vanguard Study, which aimed to pave the way for the Main Study, was developed to evaluate study procedures and recruitment methods that could be implemented in the Main Study. Launched in 2009, the Vanguard Study managed to collect data from approximately 6,000 families. By 2014, the National Institutes of Health (NIH) made the decision to cancel the Main Study and terminate data collection for the Vanguard Study (NCS, 2016). This decision came after an expert review group, the NCS Working Group, concluded that the planned NCS was far too complex, failed to incorporate innovative technology (e.g. omics), and was ultimately hindered by too burdensome a management process (McCarthy, 2014). For example, the planned study proposed 31 data collection sweeps for every participant, consisting of 17 in-person home visits (with approximately 300 variables per visit) and 14 remote data collections (Hirschfeld, 2018). Ultimately, the NCS Working Group agreed that although "the overall goals and intent are meritorious and should be a priority for future scientific support, the NCS, as currently outlined, is not feasible" (Altman *et al.*, 2014, p.5).

While the NCS Main Study was terminated before its planned 2015 launch, the preceding Vanguard Study unveiled important lessons for best practice in recruiting a nationally representative probability sample of mother–infant dyads. Pilot studies revealed that, when recruiting during pregnancy, collaborating with obstetric providers is more effective in establishing a larger and more representative sample when compared with traditional household-based recruitment tactics. Although provider-based recruitment strategies were the most successful means of obtaining a high percentage of enrolments during the first trimester, this method was not sufficient to recruit the planned pre-conceptional cohort of women (Hudak *et al.*, 2016).

### 3.2.7.2 The UK Life Study

Set against the success of the MCS, the UK Life Study proposed an innovative design and planned to collect data on 80,000 British babies born between 2014 and 2018. The multidisciplinary nature of this study combined core topics such as the influence of inequality, diversity and social mobility on child outcomes, education and school readiness, health and well-being, early life determinants of diseases, and socio-emotional and behavioural development, alongside the influence of neighbourhoods and the social environment (Davis-Kean *et al.*, 2018. The Life Study's first dedicated recruitment centre opened in January 2015, with the goal of recruiting up to 16,000 prospective mothers (of a total target of 80,000) into the cohort. Unfortunately, between January and early September 2015, a mere 249 women signed up (Pearson, 2015). Specific recruitment challenges experienced at the first Life Study Centre included:

> Midwife shortages and difficulties in their recruitment impacting on contact with and recruitment of mothers initially; difficulties in re-contacting mothers to confirm appointments following initial contact with them in scan clinics; participant burden related to questionnaire length and complex consent for three contacts as well as multiple record linkages; and challenges in engaging deprived and difficult-to-reach communities. (Dezateux *et al.*, 2016, p. 9)

On top of this, additional challenges arose during the pilot stage while testing planned recruitment procedures using NHS records and birth registrations. The percentage of mothers opting in to further contact via these administrative records was remarkably low and considered unlikely to yield a representative national probability sample (Davis-Kean *et al.*, 2018; Dezateux *et al.*, 2016). Subsequently, a review of the study by the Economic and Social Research Council (ESRC) and the Medical Research Council (MRC) identified serious challenges in recruitment as a reason to discontinue government funding for the Life Study (Pearson, 2015; UCL, 2016). Given the poor uptake levels seen at this initial Life Study Centre, the decision was taken to cancel the Life Study in early 2016 – less than 1 year after the cancellation of the NCS in the USA (Davis-Kean *et al.*, 2018.

A core parallel can be drawn between the two cancelled birth cohort studies: namely the ambition to answer highly diverse research questions (e.g. the influence of social inequalities; early origins of disease) which placed subsequent strain on the studies' designs (Pearson, 2015). For instance, broad, detailed questionnaire content is often a key prerequisite for high-quality longitudinal studies within the field of social sciences, whereas more epidemiologically focused studies often emphasise addressing specific causal questions via biomedical sampling. Concurrent with these principles are varied viewpoints and demands on cohort representativeness, sample size calculations, and study design (Davis-Kean *et al.*, 2018). Corresponding to expert advice that a future UK birth cohort study should be less complicated and ambitious than the cancelled Life Study, Davis-Kean and colleagues (2018) warn that "combining both research models [epidemiology and social science] without a clear distribution of priorities and responsibilities has the potential for resulting in overburdening study content and sampling design" (p. 24).

In fact, a core reason for recent declines in birth cohort recruitment rates is the subjective experience of being too busy by potential research participants (Vercruyssen *et al.*, 2014). The planned burden and intrusiveness of a birth cohort study necessitates careful consideration and refinement, in addition to clear communication to participating families at recruitment about what exactly will be required of them (Smyth, 2009). To help boost recruitment among busy participants, future birth cohort studies should actively strive to reduce the time and effort required by relying on modern advances in data collection (e.g. online surveys, linkage with administrative databases) (van Gelder *et al.*, 2020).

Sullivan, Joshi and Williams (2020) propose that the main lesson emerging from the cancellation of the NCS and the Life Study is the need to balance scientific ambition against practical feasibility and respondent burden. To accommodate the increasing demands on people's attention nowadays, prioritising methods to support recruitment and retention (e.g. ongoing contact between waves, birthday cards, incentives, gifts, inviting participants to presentations/events) will be essential to the success of any new birth cohort study (Sullivan, Joshi & Williams, 2020). For example, the design of a future birth cohort study in Ireland might involve keeping the study instruments quite minimal (with the potential to add greater complexity at a later stage) in order to help engage mothers and families at the outset and enhance their willingness to commit to the study (Sullivan, Joshi & Williams, 2020). Another promising solution might be a hybrid approach that integrates linked data, participant self-reports and biomedical samples into a single holistic, low-burden longitudinal framework that also regularly links with health registries (Sung *et al.*, 2021). However, there are substantial challenges associated with the standardisation, coverage and quality of routinely collected data in Ireland; thus, for such a hybrid approach to work, an in-depth evaluation of potential data sources must be conducted using the relevant statistical and technical criteria (Greene *et al.*, 2001). Planning to take full advantage of available administrative datasets, GenV, currently in its recruitment phase, intends to create a whole-of-state birth cohort study in Victoria, Australia. By design, GenV will incorporate linkage with existing records and services data, biosamples (including from pregnancy), and minimal check-ins for participant-reported information that is not routinely collected; all of this will cover pre-morbid to long-term outcomes. Learning from the preceding Life Study and NCS cohorts, GenV has been designed to boost uptake and retention across the life span in order to establish a more holistic understanding of childhood while placing a minimal burden on participating families (GenV, 2022b; Sung *et al.*, 2021).

### 3.2.8 Data protection

Distinct ethical concerns arise concerning child welfare and data protection. Legal responsibility to uphold children's and young people's rights stems from their legal status, their awareness and experience of the world, their varying levels of understanding, and their limited independence and autonomy – each of which needs special attention in order to guarantee an appropriate, ethical research process

(DCYA, 2012). All research endeavours that involve the processing of children's personal data must be capable of demonstrating compliance with both legal requirements and ethical principles (NREC, 2021).

In Ireland, the GDPR and the Data Protection Act 2018 (which gives further effect to the GDPR within Irish law) legislate a broad range of research-related activities, including data collection, data storage, data use and the disclosure of personal data (DPC, 2020). The GDPR came into effect in 2018, with the overarching goal of safeguarding all EU citizens from privacy and data breaches in today's data-driven society (Clarke *et al.*, 2019). To ensure that the processing of personal data is fair and legal under this new legislation, all persons (both children and adults) must be supplied with sufficient information about what their data will be used for, with whom it will be shared, and for how long it will be stored (UCL, 2019).

For the first time in EU data protection law, the GDPR emphasised the unique circumstances and risks (e.g. personal data breaches, profiling) faced by children when their personal data are gathered and processed without ample safeguards (DPC, 2021). Prior to the GDPR, there was no reference to children under the old EU data protection law. Now, placing children at the forefront of Europe's data protection landscape, legislation which gives effect to the GDPR stresses the importance of clear communication with children and acknowledges that, dependent on their age and level of maturity, children may be less aware of the risks associated with the processing of their personal data (DPC, 2021). For instance, organisations have an express obligation under the GDPR to verify that any information about data processing directed at child participants should be written in transparent and simple language so as to ensure the child's comprehension of it (Article 12.1, GDPR) (DPC, 2019). Moreover, the GDPR outlines the right of all children to exercise their data protection rights. For instance, all children have the right to request the erasure of their personal data; this prevents them from being inconvenienced in later life by choices they made when they might have had less awareness of the consequences of sharing their personal data (DPC, 2020). However, the Irish Data Protection Commission (DPC) also recognises that, although the GDPR gives both adults and children the rights of access, erasure, and restriction of processing, certain ambiguity remains concerning when/how children's rights should be exercised independently of their parents, and with the child's best interests in mind (DPC, 2020). When developing a future birth cohort study under new data protection legislation, GDPR compliance essentially begins with GDPR understanding and an awareness of participants' rights. The study team must select the suitable, lawful basis for data processing activities (Article 6, GDPR); and recognise the legal obligations and principles which are embedded in the GDPR and relevant national data protection legislation, such as those pertaining to personal data processing (Clarke *et al.*, 2019).

Moreover, to ensure child welfare and protection, all research is required to be conducted in accordance with *Children First: National Guidance for the Protection and Welfare of Children* (DCYA, 2017), published by the Department of Children and Youth Affairs (since renamed the Department of Children, Equality, Disability, Integration and Youth). The Children First Act 2015 informs the child welfare and protection system in Ireland by ensuring that Tusla (the child and family agency) is notified about all child protection concerns without delay (DCYA, 2017). In line with Children First, research organisations must develop a child protection policy, and should appoint a designated liaison person to be responsible for the implementation of this policy (Tusla, 2021). Under the Children First Act 2015, research organisations are also required to carry out a risk assessment to identify whether children could be harmed while participating in the study, to develop a child safeguarding statement that details the specific policies and protocols in position to mitigate identified risks, and to ensure that all research personnel are Garda vetted and have undergone the relevant child welfare and protection training (DCYA, 2017). Furthermore, although ensuring confidentiality is critical to research practice, a limitation exists in research with children wherein confidentiality must be broken if a child protection concern arises. The limits of confidentiality should be clearly outlined when obtaining informed consent (Tusla, 2021).

### 3.2.9 Potential sampling frame in Ireland

A sampling frame is the list of potential participants from which the study sample is drawn; therefore, the quality of the sampling frame influences the quality of the sample (Mohadjer *et al.*, 2016). For example, when identifying the sampling frame for Cohort '08, it was concluded that:

> The ideal sampling frame for this statistical survey is an up-to-date and fully comprehensive listing of all nine-month-olds in the country. Each infant should appear once and once only; there should be no omissions and no duplication. In addition, the frame should not include any infants who were not validly in the population, e.g. those outside the age range. (Thornton *et al.*, 2013, p. 18)

Therefore, each child aged 9 months would appear in the frame no more than once, with a calculable selection probability (Thornton *et al.*, 2013). After deliberation, the Child Benefit Register was selected by the GUI study team as it came very close to meeting each of these requirements (see Section 3.3 Case studies for further detail).

In a recent report commissioned to review theoretical sampling design options for a new UK birth cohort study (ELC-FS), Sullivan and colleagues (2020) suggest that a combination of Office for National Statistics (ONS) birth registration records and NHS birth notifications could form an optimal sampling frame of infants across the UK. Taken together, these administrative databases include information regarding home address at registration, postcode, name of the mother and infant, the NHS number of each, and whether the birth was live. Alongside constructing the original sampling frame, these administrative birth records could be useful for supporting initial non-response analyses. However, the researchers acknowledge that the actual feasibility of combining birth registration and NHS notifications is subject to extensive discussion with the NHS and ONS and must be informed by legal advice concerning the use of these sampling frames under the GDPR and other general data protection regulations. The applicability of this proposed sampling frame will be contingent on a variety of complex factors, such as (perceived) public acceptability, adherence to the GDPR and other general data protection legislation, and the actual administrative feasibility of using ONS and NHS databases for sampling purposes (Sullivan, Joshi & Williams, 2020). These potential barriers are not confined to the UK context; in fact, access obstacles to population registers and administrative databases are a critical issue recognised by researchers across many European countries (Harron *et al.*, 2017; Scherpenzeel *et al.*, 2016). In addition to data protection constraints, there are many practical issues, such as diversity in access procedures, search facilities, fees, and online access to administrative databases (Scherpenzeel *et al.*, 2016). Such barriers are particularly pertinent to the Irish context, wherein the country's health and social care data infrastructure is less centralised and relatively inaccessible in comparison to its European peers (see Chapter 2, Section 2.2.6 for further detail).

Although concrete recommendations for a sampling frame are beyond the scope of this review, a brief outline of the strengths and limitations associated with four potential sampling frames within the Irish context is shown in Table 3.1. These suggestions should be regarded as preliminary to future feasibility studies and would warrant legal and data protection expertise.

**Table 3.1** Potential sampling frames within the Irish context

| Sampling frame | Description | Implications |
|---|---|---|
| **Child Benefit Register** | The Child Benefit is a monthly payment to the primary caregiver of children aged under 16 years living in the Republic of Ireland. Children aged 16 and 17 years can continue to receive the Child Benefit if they remain in full-time education or full-time training, or have a disability. Once the birth of a child has been registered, the parent/guardian will be automatically contacted to register for the Child Benefit (Citizens Information, 2021a). | ● Financially advantageous for parents to register their child. Very few parents in Ireland do not claim Child Benefit.<br>● Child Benefit is a universal payment; all parents in Ireland are eligible. Income level and Pay Related Social Insurance (PRSI) payments do not influence eligibility.<br>● Each member of the target population appears just once on this centralised register.<br>● Important information is available, such as contact information and social security number.<br>● Available in an easily accessible, electronic format.<br>● The Child Benefit Register was previously successful in providing the sampling frame from which the GUI Infant Cohort was drawn.<br>● However, as the sampling procedure used to recruit Cohort '08 predates the GDPR and other general data |

| Sampling frame | Description | Implications |
|---|---|---|
| | | protection regulations, accessing sociodemographic data and contact details from the Child Benefit Register might no longer be permitted today.<br>● Legal advice is therefore needed in order to determine whether this register could still be used as a potential sampling frame under modern data protection laws. |
| **Birth Notification Form (Form BNF/01**) | A Birth Notification Form (Form BNF/01) is typically completed with the parent(s) by hospital staff (for hospital births), or by a doctor or midwife (for home births), to ensure that accurate information is recorded. Form BNF/01 contains the information to be recorded on the child's birth certificate (once registered within 3 months of the birth). The form is then forwarded to the Registrar of Births' office to inform the Registrar that a birth has occurred. The information recorded on this form includes the infant's surname and forename, the time, date and place of birth, and sex of the infant, alongside details about the mother and father (forename(s), surname(s), date of birth, home address, PPSN, occupation, nationality, civil status) (Citizens Information, 2021b). | ● This form is filled out for all infants born in Ireland (nationally representative sampling frame).<br>● Contains sufficient detail to facilitate initial contact with potential participants (e.g. home address) and subsequent tracing between waves (e.g. PPSN).<br>● Obtaining access to these secure records would require consent from the HSE Civil Registration Service and/or individual maternity units/patients.<br>● Accessing these data may not be possible under the GDPR.<br>● This sampling frame would not include newly arrived infants (e.g. those resident in Ireland at age 9 months but who were born outside the State). |
| **Population census** | The census is a count and record of all persons in Ireland that occurs every 5 years. The most recent census occurred on Sunday 3 April 2022. The CSO protects all census data; it is secure, GDPR compliant and protected by law. The CSO only releases grouped, anonymous data in the form of statistics (e.g. population distribution, population age, marital status, sex, place of birth, occupation, and religion). All completed census forms are securely stored for 100 years, after which they become | ● Incomplete and inaccessible.<br>● No contact information available for sampling purposes.<br>● While census data are secure and inaccessible, group statistics could help inform sample stratification (e.g. for minority groups).<br>● Census data are only collected every 5 years, and so may not reflect recent trends (e.g. migration). |

| Sampling frame | Description | Implications |
|---|---|---|
| | publicly accessible, in accordance with the Statistics Act 1993 (CSO, 2021). | |
| **The Maternal and Newborn Clinical Management System**<br>**(MN-CMS)** | The MN-CMS Project refers to the ongoing implementation of a new electronic health record (EHR) for all women and infants being cared for in maternity, newborn, and gynaecology services in Ireland. This database helps all maternal, newborn, and gynaecology information to be shared with appropriate providers of care in accordance with the GDPR. The MN-CMS was successfully implemented at Cork University Maternity Hospital in December 2016, at University Hospital Kerry in March 2017, the Rotunda Hospital in November 2017, and at the National Maternity Hospital in January 2018. The MN-CMS will be deployed to the remaining 15 maternity hospitals and units in Ireland on a phased basis (eHealth Ireland, 2021). | • Easy to access, single record on mother and infant that is electronically accessible across all maternity hospitals.<br>• Full roll-out of this system is ongoing, and may not be implemented for another few years.<br>• In its current phase, this system would limit the sample to only 4 out of the 19 maternity units/hospitals in Ireland – i.e. would not be nationally representative.<br>• Obtaining access to this system for research purposes might not comply with the GDPR.<br>• Prior consent would be required to access the contact details of patients for sampling and recruitment purposes. |

## 3.3 Case studies

### 3.3.1 Growing Up in Ireland

#### 3.3.1.1 Cohort age

The overall GUI study consists of two separate cohorts: Cohort '08 (recruited at 9 months old) and Cohort '98 (recruited at 9 years old). This parallel cohort design meant that policy-relevant data concerning school-aged children would be available much earlier than if data collection began only in infancy (Sullivan, Joshi & Williams, 2020). This report will focus on Cohort '08 (formerly known as the Infant Cohort).

#### 3.3.1.2 Sampling frame and recruitment

The objective of Cohort '08 was to generate and interview a representative sample of 9-month-old children living in the Republic of Ireland. The Child Benefit Register was identified as a useful sampling frame. The Child Benefit Register is a payments database maintained by the Department of Social Protection, which is responsible for ensuring that it is current and kept fully up to date. The Child Benefit must be claimed within 6 months of the child's birth (or 6 months since moving to Ireland). Considering the Child Benefit Register is used to facilitate monthly payments in the interests of all children aged under 16 years, it is financially beneficial for parents to ensure that their child is included on the register (McNamara, Murray & Williams, 2019). Omissions from this register are extremely uncommon; the Department of Social Protection conducts periodic postal checks of recipients and follows up on non-contacts from those postal checks (Quail *et al.*, 2011a). Hence, the Child Benefit Register had the following optimal characteristics required for use as a sampling frame: the database comprises an up-to-date directory of eligible members of the study population, contains a broad range of relevant characteristic variables (e.g. PPSN, family contact details, nationality, child's date of birth, and mother's marital status), and already exists within an easily accessible electronic format (Quail *et al.*, 2011a). In fact, data from the Child Benefit Register were validated against the CSO Live Births Registration, with records across the two independent databases found to be extremely consistent (McNamara, Murray & Williams, 2019).

#### 3.3.1.3 Sample size calculation

The 2001 GUI study design brief (Greene *et al.*, 2001) determined the target sample size as:

> A high-quality probability sample size which will yield an achieved sample of not less than 10,000 children, one fifth of annual births, in the birth cohort and not less than 8,000 children in the older cohort. Recent experience with increasing non-response rates in Irish surveys (a trend common to many developed countries) suggests that the initial sample selected might have to be 20 to 30 per cent higher to attain these targets. (Greene *et al.*, 2001, pp. 16–17)

Rather than a traditional cohort design (i.e. a single cohort recruited at a given time point and revisited at fixed intervals), Greene *et al.* (2001) also proposed that a parallel cohort design would enable comparison across time and sociocultural context once the infant cohort effectively 'catches up' with the child cohort.

From the Child Benefit Register, eligible participants were identified as those who would be aged 9 months at the planned time of fieldwork (September 2008 to April 2009). Thus, the date of birth of eligible children lay between 1 December 2007 and 30 June 2008. This yielded a total eligible population of 41,185 infants born during the specified time period (Quail *et al.*, 2011b). Of those 41,185 eligible for the cohort, a sample was chosen on a systematic basis, pre-stratifying by the payee's (i.e. mother's) marital status, nationality, and county of residence, alongside the number of children in the family for which the Child Benefit was claimed (all of these variables were accessible internally from the information documented on the Child Benefit Register) (Quail *et al.*, 2011a). A simple systematic selection method based on a random start and constant sampling fraction was employed (McNamara, Murray & Williams, 2019).

Initial contact was made by mailing detailed information letters to all potential participants which outlined the type of data that would be collected, the longitudinal nature of the GUI study and the voluntary nature of the research (Thornton *et al.*, 2013). The interviewer subsequently made an in-person visit to each household to answer any questions, obtain signed consent forms and arrange an interview with the primary and secondary caregiver (if applicable) (Swift *et al.*, 2021). As a result of this selection process, 11,134 infants and their families were recruited into Wave 1 of the GUI Cohort '08 and partook in household interviews during the period September 2008 to April 2009 (Quail *et al.*, 2011b). The complete sample size represented just above one-quarter (27%) of all births in Ireland during the field data collection period (Thornton *et al.*, 2013). The response rate represents approximately 65% of all families approached, and 69% of valid contacts made throughout the course of fieldwork (Williams *et al.*, 2010).

#### 3.3.1.4 Ensuring representativeness

Corresponding to best practice, the complete sample of 11,134 was statistically reweighted according to external population estimates to confirm whether it was fully representative of all infants in Ireland (Williams *et al.*, 2010). The variables and population distributions used for weighting were drawn from two sources: (1) the number and characteristics of infants (aged under 1 year) and their families as per the 2006 Census of Population, and (2) the Child Benefit Register from which the sample was drawn (Thornton *et al.*, 2013). The 73,662 infants recorded on the Child Benefit Register for the 2008 calendar year were selected as the population against which the sample was statistically weighted and grossed in statistically re-adjusting the sample. The system used to generate the sample weights was based on a minimum information loss algorithm which made sure that the distribution of cases in the completed GUI sample fit with a series of control totals for the general population. The infant was the unit used in this weighting system and a total of 11 primary characteristics of the infant and their family were applied in the generation of the weights: family structure, mother's age, mother's principal economic status (PES), father's PES, family's social class, mother's education, household tenure, region/child's gender (i.e. separate geographical categories by gender used to ascertain geographical representativeness, with regions ranging from border counties to the west), mother's marital status, mother's nationality, and mother's residency status (Quail *et al.*, 2011a).

In addition to reweighting the complete sample according to known population figures, response rates from the pilot and dress rehearsal stages of the GUI study highlighted the necessity to oversample smaller (but policy-relevant) subgroups of the population to guarantee a suitable absolute number for analysis. For instance, pilot response rates were lower among families wherein the mother's marital status (as per the Child Benefit Register) was something other than 'married'. Accordingly, those from each of the non-married categories were slightly oversampled in the completed sample, whereas those in the 'married' group were undersampled (by 6.1 percentage points) (Thornton *et al.*, 2013). Moreover, to ensure that non-national infants were adequately represented in Cohort '08, an independent supplementary sample of 700 non-national infants was included (after the main selection process) to address the higher levels of non-participation observed among non-national families. Consequently, the inclusion of this supplementary sample within the completed sample confirmed that the total number of non-national participants in the final sample for analysis was representative of external population estimates (Quail *et al.*, 2011a). In addition to English and Irish, information sheets and questionnaires were made available in Romanian, Latvian, Lithuanian, Mandarin Chinese, French and

Polish. Upon request, a translator could also be provided for the household visit (Thornton *et al.*, 2013).

### 3.3.1.5 Timing of additional waves

The target sample for all additional waves is comprised of the 11,134 who participated at Wave 1, apart from those who left Ireland, those who passed away between waves, and those who withdrew from the study ('hard refusals'). Those who did not respond to any one wave due to time constraints or other reasons ('soft refusals') were still included in the sample at the subsequent waves (Quail *et al.*, 2019). As seen in Table 3.2, follow-up waves are typically conducted with Cohort '08 every 2–3 years. A supplementary online COVID-19 questionnaire was first administered in December 2020, and the cohort was revisited again in 2021–2022 when they were aged 13 years.

### 3.3.1.6 Retention process

Cohort '08 had attrition at each wave of approximately 10–12% (see Table 3.2). To address this attrition at each round of interviewing, data were reweighted prior to any analysis by means of an iterative procedure based on core variables assessed at previous waves (e.g. maternal age, family structure, income). The longitudinal weights generated at each wave help to tackle potential biases (such as selective attrition) so as to ensure ongoing representativeness (McNamara *et al.*, 2020).

Considering the significant cost of birth cohort studies such as the GUI study, efficient retention strategies that engage and retain cohort participants are imperative to the integrity of research outcomes (Teague *et al.*, 2018). Given the heavy reliance on face-to-face interviewing for most rounds of the GUI study, interviewer training focused on the utmost importance of establishing a good rapport with respondents and gaining their trust prior commencing the formal interview procedure (Thornton *et al.*, 2013). The use of reminders was another pertinent means of retaining participants. For instance, to counteract low response rates often associated with postal surveys, a telephone reminder was included during Wave 4 data collection to help boost response rates (MacNamara Murray & Williams, 2019). In addition, small gifts were offered to acknowledge participation in the GUI study. For instance, 3-year-old children who completed Wave 2 of data collection were given a wall chart to measure their height alongside a colouring book and pencils, whereas 9-year-old children received a Growing Up in Ireland-branded crayon and notebook gift set (MacNamara, Murray & Williams, 2019; McCrory *et al.*, 2013).

**Table 3.2** Timing of additional GUI data collection waves and response rates

| Data collection | Year | Child age | Participants | Response rate |
|---|---|---|---|---|
| **Wave 1** | 2008–2009 | 9 months | 11,134 | |
| **Wave 2** | 2010–2011 | 3 years | 9,793 | 88.0% |
| **Wave 3** | 2013–2014 | 5 years | 9,001 | 80.8% |
| **Wave 4** (interwave postal survey) | 2015–2016 | 7/8 years | 5,344 | 48.0% |
| **Wave 5** | 2017–2018 | 9 years | 8,032 | 72.1% |

**Note:** Response rate is expressed as a percentage of the total number of participants ever interviewed (N=11134).
Source: Quail *et al.*, 2019

## 3.3.2 Danish National Birth Cohort

### 3.3.2.1 Cohort age

The Danish National Birth Cohort (DNBC) is a nationwide birth cohort study that, during the period 1996 to 2002, recruited more than 100,000 pregnant women residing in Denmark (Olsen *et al.*, 2001). While birth cohort studies typically focus exclusively on child outcomes, the DNBC began recruitment during early pregnancy to facilitate comprehensive data collection throughout gestation and into early motherhood. Since the inception of the DNBC, the recruited mothers and their children have participated in frequent prospective follow-up waves throughout childhood (Bliddal *et al.*, 2018).

### 3.3.2.2 Sampling frame and recruitment

The target population consisted of all clinically recognised pregnancies among Danish-speaking women who wished to carry their pregnancy to term and were living in Denmark at the time of the recruitment (Jacobsen, Nohr & Frydenberg, 2010). Although the initial goal of the DNBC was to recruit women as early as possible in their pregnancy, the National Board of Health (now called the Danish Health Authority) did not permit the study team to contact pregnant women directly for recruitment purposes. Alternatively, collaboration was sought with Danish GPs, of whom there are approximately 3,500 in the country (Olsen *et al.*, 2001). GPs were deemed a suitable means of recruitment as they would meet almost all eligible participants at least once during pregnancy and were usually their initial point of contact with the antenatal healthcare system (Jacobsen, Nohr & Frydenberg, 2010). Accordingly, when women came for their first pregnancy visit, provided by GPs at around 6–12 weeks' gestation, all participating GPs offered a consent form and background information about the DNBC, also referred to as 'Better Health for the Mother and Child' (Olsen, 2012). There was also a backup recruitment procedure in place for women who did not receive an invitation at the GP: expectant mothers were invited by the midwife whom they were referred to for follow-up after their initial GP visit (Olsen *et al.*, 2001).

Yet, it is important to acknowledge that a substantial proportion of Danish GPs did not agree to take part in the recruitment process; thus, around 40% of eligible women were never given an invitation to participate in the study (Jacobsen, Nohr & Frydenberg, 2010). Of the pregnant women who did receive an invitation from participating GPs, an estimated 60% agreed to participate by returning a signed consent form (Morales-Suárez-Varela *et al.*, 2018). However, this is merely an estimated percentage; the actual participation rate among pregnant women is unknown because participating GPs did not agree to report their own participation in the recruitment process (Jacobsen, Nohr & Frydenberg, 2010).

### 3.3.2.3 Sample size

Regarding power calculations, the DNBC relied on present incidence rates of rare childhood diseases (e.g. congenital malformations, genital malformations, facial clefts, childhood cancers, leukaemia) to determine the large sample size required to successfully examine rare exposures and outcomes, and to facilitate the potential exploration of gene–environment interactions. This required sample size was calculated using the smallest detectable relative risk (RR) in a case-control analysis nested within the cohort, using four controls per case (Olsen *et al.*, 2001). Consequently, to achieve adequate statistical power, 91,389 women across Denmark with 100,421 pregnancies (including twins, triplets, etc.) were recruited into the cohort from 1996 to 2002. This complete sample corresponds to approximately 30% of all pregnancies that occurred in Denmark throughout the DNBC recruitment period (Bliddal *et al.*, 2018).

### 3.3.2.4 Representativeness

From its outset, the DNBC prioritised obtaining a large sample size ahead of recruiting a sample that is entirely representative of the overall Danish population. Eligible participants included all pregnant women living in Denmark who wished to carry their pregnancy to term and spoke Danish sufficiently well to participate in the telephone interviews; no further exclusion criteria were used. Therefore, it was never anticipated that the resulting cohort would provide a representative sample of pregnant women. Instead, the intent was to use the large mother–child cohort for internal comparison of those exposed and unexposed to certain risk factors occurring in early life. The rationale was that selection biases should not be associated with both the exposure (e.g. certain environmental pollutants) and outcome (e.g. infection) since the outcome is unknown at the time of recruitment (Olsen *et al.*, 2001).

While the purpose of the DNBC was never to generate data for descriptive studies of representative distributions of the Danish population, the non-participation rate (approximately 40% of all women invited) remained a cause for some concern regarding the validity of the cohort (Jacobsen, Nohr & Frydenberg, 2010). Following an analysis of initial participation at recruitment, Greene *et al.* (2011) observed that, while the women who chose to participate in the DNBC were generally of higher social status, were healthier and had lower levels of disease than all those eligible for the cohort, differential selection was modest and the impact of selection bias on multiple exposure–outcome associations was limited. These findings indicate the potential merit of recruiting a subset of motivated participants willing to engage in long-term follow-up versus prioritising representativeness when designing epidemiologically focused birth cohort studies (Greene *et al.*, 2011).

### 3.3.2.5 Timing of additional waves

Following on from the four initial CATIs conducted with women during pregnancy and after the child's birth (see Section 2.3 Case studies), the DNBC has extended data collection with additional rounds occurring throughout infancy, childhood and into adolescence (see Table 3.3).

### 3.3.2.6 Retention process

With the passage of time between data collection waves (see Table 3.3), the motivation to participate in longitudinal research can vary due to changes in life situations, social or health conditions, and lifestyle (Bliddal *et al.*, 2018). Whereas the large sample size of the DNBC is advantageous in providing greater statistical power than smaller birth cohorts, the substantial number of participants limits the implementation of strategies to minimise attrition, such as having ongoing contact with cohort members and providing incentives (e.g. book tokens, birthday cards). For instance, the expense of sending one single postcard to all DNBC children was estimated to be more than €150,000 (Lawlor & Mishra, 2009). Fortunately, in the case of loss to follow-up, consent for extensive linkage to nationwide health databases (provided by participating mothers at recruitment) allows for follow-up information to be extracted regarding disease and death among those who stopped responding in follow-up waves (Nohr & Liew, 2018). Participants who do not participate in one or more waves (i.e. 'soft refusals') are not withdrawn from the cohort permanently; instead, they are re-invited to participate in subsequent waves. As of 2011, less than 0.5% of mothers have formally withdrawn consent to participate in all future DNBC data collections (Greene *et al.*, 2011). Recent data from the maternal follow-up (2013–2014) found that of the mothers who participated in the first pregnancy interview, 90.3% (n=82569) were eligible for this sub-study; just 5.6% of mothers were not eligible due to emigration or withdrawal of consent, and a small number of mothers (n=449) were no longer eligible due to death (Bliddal *et al.*, 2018).

**Table 3.3** Timing of additional DNBC data collection waves and response rates

| Data collection wave | Invited | Timing | Respondents | % of invited | % of overall sample |
|---|---|---|---|---|---|
| Enrolment in the DNBC | 100,421 | 1996–2002 | | | |
| Interview 1 (prenatal) | 100,421 | 12 weeks' gestation | 92,892 | 92% | 92% |
| Interview 2 (prenatal) | 100,421 | 30 weeks' gestation | 87,802 | 87% | 87% |
| Food Frequency Questionnaire | 100,421 | 25 weeks' gestation | 72,821 | 72% | 72% |
| Interview 3 (postnatal) | 100,421 | When the child was aged 6 months | 70,292 | 69% | 69% |
| Interview 4 (postnatal) | 100,421 | When the child was aged 18 months | 66,764 | 65% | 65% |
| 7-year follow-up | 91,256 | When the child was aged 7 years (November 2005 to July 2010) | 57,282 | 63% | 57% |
| 11-year follow-up | 90,986 | When the child was aged 11 years (July 2010 to August 2014) | 49,963 | 55% | 50% |
| Maternal follow-up | 78,010 | Mothers were invited to complete follow-up measures from 2013 to 2014. | 43,641 | 55% | 43% |
| 14-year follow-up | 78,651 | Measures on the dietary habits of 14-year-olds began in April 2013 and ended in 2017. | 36,599 | 47% | 36% |
| Puberty follow-up | 22,439 | Puberty follow-up began in 2012 and ended in February 2021. This was a sub-cohort within the DNBC of those born between 2000 and 2003. | 15,819 | 70% | 16% |
| DNBC COVID-19 study | 53,323 | Seven weekly data collections about the impact of COVID-19 (March to May 2020) | 25,898 | 46% | 26% |

***Note:*** 'Invited' refers to those who are contacted to participate at each wave. This figure excludes those ineligible to participate due to death, emigration, or withdrawal of consent to participate in future data collections. Additional eligibility criteria for specific waves also determined who could be invited (e.g. only DNBC participants who had consented to provide their email address and/or phone number could participate in the COVID-19 data collection; only those born from 2000 to 2003 were invited to take part in the puberty sub-cohort).

Source: DNBC, 2020c

## 3.3.3 Étude Longitudinale Française depuis l’Enfance

### 3.3.3.1 Cohort age

Étude Longitudinale Française depuis l’Enfance (ELFE), a multidisciplinary nationwide birth cohort, involves 18,329 infants born in 2011 across a random sample of 349 maternity wards situated in mainland France (Camier *et al.*, 2020). The initial goal of the ELFE cohort was to initiate data collection during pregnancy in order to accurately describe the earliest stages of child development, with a particular focus on children’s physical and social environment (Pirus *et al.*, 2010). Yet, it was established by the ELFE research team that France’s statistical and healthcare system would be incapable of forming a nationally representative sample of pregnant women who could be monitored at a suitably early stage in their pregnancy. Instead, consistent with the pre-existing model of France’s national perinatal surveys (Enquêtes Nationales Périnatales; ENP), recruitment would occur in maternity units in the days immediately following the child’s birth (Pirus *et al.*, 2010). This approach makes ELFE the first national birth cohort to exist in France in which the determinants of children’s development, health and socialisation are studied from birth to adulthood (Charles *et al.*, 2020).

### 3.3.3.2 Sampling frame

To ensure representativeness, the objective was to draw a sample from all of the women giving birth in metropolitan France (including mainland France and Corsica) on days predetermined by the National Institute of Statistics and Economic Studies (Institut national de la statistique et des études économiques; INSEE) (Pirus *et al.*, 2010). However, since the sampling procedure necessitated personal contact with mothers in the maternity hospital, obtaining a sizeable representative sample of births would not be a straightforward task. Most mothers only spend a short duration in hospital following childbirth (on average, 2 or 3 days), and infants are delivered in 540 different public and private maternity units located across mainland France. Therefore, sampling proceeded in four consecutive waves across a 1-year period (to account for potential seasonal variations) and the decision was made to reduce the number of maternity units to 349 (Charles *et al.*, 2011). A stratified sampling approach that accounted for the size of each maternity unit was then adopted to facilitate oversampling in larger units, thus decreasing data collection costs (Charles *et al.*, 2020).

Of the 349 selected units, infants were recruited from the 320 maternity units that consented to take part in the ELFE study. The recruitment period took place across 25 days in 2011, grouped into four intervals of 4–8 days throughout the year (Charles *et al.*, 2020). These four periods included 1–4 April, 27 June to 4 July, 27 September to 4 October, and 28 November to 5 December (ELFE, n.d.b). Within this recruitment period, 12 days were selected to coincide with the Permanent Demographic Sample of the INSEE, a routine national survey of all children born on particular days each year. Accordingly, for almost one-half of the cohort, participating ELFE children could be compared with children of the general population born across the same 12-day period (Duşa *et al.*, 2014). This position of ELFE as sub-sample of the Permanent Demographic Sample facilitates tracking of the cohort's representativeness as attrition increases across waves (Pirus *et al.*, 2010).

#### 3.3.3.3 Sample size calculation

A priori calculations indicated that, from the total sample of 349 maternity units, around 40,000 births would take place during the ELFE recruitment period. Estimating a non-participation rate of up to 10% at maternity unit level, and 45% at the individual level (as per figures observed in pilot studies), the chosen sample of maternity units should be adequate to recruit around 20,000 infants (Pirus *et al.*, 2010). A sample size of approximately 20,000 births (2.5% of annual births) was deemed achievable while also providing the statistical power necessary to address ELFE's primary research questions (Charles *et al.*, 2011).

#### 3.3.3.4 Recruitment

To be eligible for the ELFE birth cohort, infants had to be born on one of the 25 recruitment days; only single or twin live births were included, and pregnancies were required to be of at least 33 weeks' duration. Participating mothers had to be aged over 18 years, capable of providing informed signed consent, and have no intention to leave mainland France within the next 3 years. Where relevant, fathers were also informed of the voluntary nature of the study and of their right to withdraw consent for their own participation. Consent forms and information letters were available in French, Arabic, Turkish and English – the languages most often spoken by women giving birth in France (Charles *et al.*, 2020).

More than 96% of new mothers (n=37494) who met these inclusion criteria were contacted by ELFE research assistants during their stay at the maternity unit, 51% of whom (n=18040) consented to take part in the birth cohort study. In total, these mothers gave birth to 18,329 infants, including 289 pairs of twins (Charles *et al.*, 2020). The final sample was slightly below the a priori sample size calculations.

#### 3.3.3.5 Timing of additional waves

As evident in Table 3.4, data collection was intensive throughout infancy and the preschool years (from 2011 to 2015), with CATIs conducted with both parents when the child was aged 2 months, 1 year, and 2 years, and again with the mother when the child was aged 3.5 years. Alongside the frequent use of CATIs, details regarding infant feeding were gathered every month from the time the child was aged 2 months to 10 months via web-based or paper-based questionnaires. With parental consent, the child's GP also received a questionnaire to be completed when the child was aged 2 years.

In addition to the main surveys administered to the entire ELFE cohort, specific measures were undertaken in subgroups (see Table 3.4). For instance, with the support of 30 district maternal and child welfare services, medical examinations were performed in 2016 at nursery schools for a subsample of 3,124 ELFE children and 6,815 of the control children born on the same days. Other supplementary follow-up measures included dust sample collection (from the child's bedroom) at birth and when they were aged 3.5 years, an arranged household visit when the child was aged 3.5 years, and a survey completed by the child's nursery school teacher when the child was aged 3–4 years (Charles *et al.*, 2020). More recently, a supplementary survey regarding experiences during the COVID-19 pandemic was completed by ELFE children in 2020 (40% participation of those contacted), and a survey with the cohort when the child was aged 10.5 years recently took place between December 2021 and August 2022 (ELFE, n.d.a).

### 3.3.3.6 Ensuring representativeness

The final sample of approximately 18,300 infants, or about 1 in 42 births in metropolitan France, was derived from a sampling frame with unequal probabilities of inclusion (Juillard, 2015). With ethical approval,[19] some information regarding the births of mothers who refused to participate in ELFE was obtained from birth certificates and stored anonymously. Gathering these data made it possible to apply a weighting system to correct for non-representativeness of the ELFE cohort against the overall French population (Charles *et al.*, 2020). The weighting procedure involved assigning a statistical weight to each of the 18,329 infants, equivalent to the number of children they represent in the target population (764,000 infants, and an approximate total of 753,500 families) (Juillard, 2015). The weights take into consideration the sampling plan, alongside refusals to participate at both maternity unit and individual levels (Charles *et al.*, 2020). After allowing for the initial weights drawn from the sampling frame, the weights were adjusted to account for non-participation at different levels: the proportion of maternity units that did not take part in the initial survey, and the proportion of mothers who gave birth on one of the relevant study days but chose not to participate. There were two types of non-participation observed for maternity units: the units that did not participate at all, and those that participated partially (i.e. non-participation on some of the selected study days). Data regarding four variables common to both participating and non-participating units (region, legal status, stratum, and level of medical authorisation) were used to compensate for the non-participation of the 29 maternity units by increasing the weighting of the 320 participating units. Furthermore, to address non-participation at an individual level, data on variables common to participating and non-participating mothers (mother’s age, district of residence, socioeconomic status, indicator of mother’s activity status during pregnancy, a primiparity indicator, a twin birth indicator, and the infant’s gestational age) were also used alongside the four variables characterising the maternity unit. Subsequently, a calibration was executed

---

[19] Ethical approval for data collection in the maternity units and for all subsequent waves was obtained from a variety of relevant committees in France: the National Advisory Committee on Information Processing in Health Research (Comité Consultatif sur le Traitement de l’Information en matière de Recherche dans le domaine de la Santé; CCTIRS), the National Data Protection Authority (Commission Nationale de l’Informatique et des Libertés; CNIL) and the committees for the protection of persons engaged in research (Comités de Protection des Personnes; CPP). ELFE was also approved by the National Council for Statistical Information (Conseil national de l’information statistique; CNIS) (Charles *et al.*, 2020).

against the civil register and the national perinatal survey (ENP), enabling the weighted sample to be matched to the target population with regard to geography and the mother's sociodemographic characteristics (Juillard, 2015).

### 3.3.3.7 Retention process

As seen in Table 3.4, participation rates for the main 2-month (92%), 1-year (86%) and 2-year (82%) parental CATIs were relatively stable, with the number of non-responders remaining quite low at approximately 4% per wave. Nevertheless, withdrawal from the study occurred across each wave of data collection, from families who gave a written/oral request to end their participation and from those who could not be contacted during any three previous consecutive waves. Although families who left metropolitan France became ineligible for telephone interviews, they were instead offered a short paper version of the questionnaire to complete (Charles *et al.*, 2020).

Charles *et al.* (2020) found that, compared to the mothers who continued to participate in data collection, mothers who had withdrawn by the end of the 3.5-year wave (n=2092) had a greater likelihood of being a single mother, being aged under 25 years, not having achieved a university degree, being unemployed, or being born outside of France. Notwithstanding these findings, no differences were observed between the two groups in terms of parity, maternal health prior to and during pregnancy, mode of delivery, and mean birthweight (Charles *et al.*, 2020).

Alongside adjusting for initial non-representativeness stemming from the sampling frame and rates of consent to participate in ELFE, weights are generated for each data collection wave to account for inter-wave attrition. Moreover, the capacity for linkage with national health insurance data, for which 95% of the cohort gave permission, enables comparisons to be made between active participants and those lost to follow-up. The medical and teacher surveys conducted within nursery schools provide another valuable method of quantifying selection and attrition biases, as they also contain data from control children not included in ELFE but born on the same day as the ELFE children (Charles *et al.*, 2020).

Considerable efforts are also made to engage with ELFE cohort members. These include regular newsletters detailing the most recent study findings, invitations to live online conferences wherein parents can ask questions to researchers about the next steps in data collection and recent study findings, and the recruitment of ELFE parent/child ambassadors who are invited to participatory workshops in which they can provide feedback on various matters and exchange ideas with the project coordinators and research team. Each of these methods aims to encourage ongoing participation and ensure that ELFE children and parents have the opportunity to play an active role in the research (ELFE, 2021).

**Table 3.4** Timing of the main ELFE surveys and response rates during the first 5 years of follow-up

| Data collection | Year | Eligible | Withdrawal | Non-response | Total participants | % of eligible participants | % of overall sample |
|---|---|---|---|---|---|---|---|
| Recruitment | 2011 | 37,494 | 70 | - | 18,040 (pregnancies) | 51% of contacted | - |
| Infant diet web/paper survey | 2011–2012 | 17,970 | 2 | 819 | 12,140 | 70% | 67% |
| 2-month parental CATI | 2011–2012 | 17.968 | 311 | 240 | 16,278 | 92% | 90% |
| 1-year parental CATI | 2012 | 17,657 | 595 | 804 | 14,436 | 86% | 80% |
| 2-year doctor questionnaire | 2013 | 17,070 | 1,090 | 468 | 7,574 | 46% | 42% |
| 2-year parental CATI | 2013 | 17,070 | 1,090 | 881 | 13,276 | 82% | 74% |
| 3.5-year parental CATI | 2014 | 16,015 | - | 1,005 | 12,032 | 80% | 67% |

***Note:***

- **Withdrawal** = total number of participants lost between waves who formally withdrew from participating in ELFE or were lost in the case of parental/child death.
- **Non-response** = those who could not be successfully contacted by the study team for each specific wave. They were recontacted again across waves.
- **Eligible** = the total number of participants who were eligible for follow-up at each wave (i.e. had not left metropolitan France, passed away, or formally withdrawn from the study).

Source: Charles *et al.*, 2020

**Table 3.5** Participation rates in data collection carried out with subgroups of the ELFE cohort from 2011 to 2016

| Data collection | Timing | Eligible | Total participants | % of eligible participants | % of overall sample |
|---|---|---|---|---|---|
| Dust collection | Birth to age 2 months (2011) | 6,390 | 3,217 | 51% | 18% |
| Home visit | 3.5 years (2014–2015) | 11,453 | 9,293 | 81% | 52% |
| Home dust sampling | 3.5 years (2014–2015) | 1,035 | 837 | 81% | 5% |
| Accelerometery | 3.5 years (2014–2015) | 595 | 463 | 78% | 3% |
| Biological sampling | At birth (2011) | 9,053 | 5,903 | 65% | 33% |
| Biological sampling | 3.5 years (2014–2015) | 3,415 | 2,125 | 62% | 12% |
| Nursery school teacher survey | 3–4 years (2014–2016) | 10,553 | 5,178 | 49% | 28% |
| Nursery school medical examination | 4–5 years (2016) | 4,458 | 3,124 | 70% | 17% |

Source: Charles *et al.*, 2020

## 3.3.4 Millennium Cohort Study

### 3.3.4.1 Cohort age

Similar to Cohort '08 in the GUI study, the MCS cohort members were first interviewed at age 9 months. Appropriately, the MCS population is characterised by a single cohort of children born between 1 September 2000 and 31 August 2001 (for England and Wales) and between 24 November 2000 and 11 January 2002 (for those living in Scotland and Northern Ireland); these children were were alive and living in the UK at age 9 months and eligible to receive Child Benefit (Connelly & Platt, 2014).

### 3.3.4.2 Sample design and representativeness

Unlike the traditional trend of former UK birth cohorts to employ a systematic random sample of all children born during a specific week, the MCS opted to sample births across a 16-month period in order to spread fieldwork across a longer, less-intensive time frame, and to detect any potential season-of-birth effects (Hansen, 2014). The MCS was also unique in being the first UK birth cohort to include all four countries of the UK. Moreover, the MCS oversampled children from deprived backgrounds in order to better understand the influence of social disadvantage on children’s outcomes (Hansen, 2014). Finally, to reflect the growing diversity of the UK at the turn of the millennium and to address the evident gaps in health, educational and social outcomes across ethnic groups, areas of relatively high ethnic minority concentration were oversampled in the MCS (Hansen, 2014). Ultimately, the decision to disproportionately represent children from socially disadvantaged and ethnic minority backgrounds was made to ensure that traditionally hard-to-reach populations were sufficiently represented, and that the resultant sample sizes were sufficient for the analysis of ethnic minorities, children from disadvantaged backgrounds, and children living within each of the four UK countries (Connelly & Platt, 2014).

The population was stratified to implement this planned sample design. First, the population in England was stratified via the stratification of electoral wards: (1) an ‘ethnic minority’ stratum wherein the proportion of minorities (according to the 1991 Census) in that ward was at least 30%; (2) a ‘disadvantaged’ stratum which included children living in wards within the poorest 25% of wards as per the Child Poverty Index (CPI) for England and Wales; and (3) an ‘advantaged’ stratum which comprised children who did not live in the ‘ethnic minority’ and/or ‘disadvantaged’ wards. Beyond England, the low proportion of ethnic minorities living in Wales, Scotland, and Northern Ireland at the time (around 1% of the population) only warranted the inclusion of two strata: (1) a ‘disadvantaged’ stratum of children who lived within the poorest 25% of wards according to the Child Poverty Index (CPI); and (2) an ‘advantaged stratum’ of all other children living in other wards in these countries (Hansen, 2014). The random selection within each stratum in each of the four countries yielded a disproportionately stratified cluster sample; families living in disadvantaged areas, for instance, would have a higher probability of selection than families living in advantaged areas. This disproportionality indicates that the sample was not self-weighting; thus, weighted estimates of means and variances were required (Smith *et al.*, 2007).

### 3.3.4.3 Sample size calculation

The initial target sample size of children at Wave 1 was 15,000 children from a population of births anticipated to be approximately 70,000 across the 16-month period (Joshi & Fitzsimons, 2016). If this sample was divided between the four UK countries according to their birth rates, the resultant sample would have been 12,600 children in England, 750 in Wales, 1,200 in Scotland and 450 in Northern Ireland. Regarding the three smaller UK countries, the sample sizes would have been insufficiently powered for statistical analysis, particularly when considering the impact of sample attrition over time. Therefore, each of the three smaller countries was assigned a sample of 1,500 children, leaving 10,500 for England (Smith *et al.*, 2007).

To ensure ongoing representativeness, the division of this targeted sample (N=15000) across the aforementioned strata was one-half in the 'advantaged' and one-half in the 'disadvantaged' electoral wards for Wales, Scotland, and Northern Ireland. In England, one-half of the sample was allocated to advantaged wards and one-quarter was allocated to each of the 'ethnic minority' and 'disadvantaged' wards (Smith *et al.*, 2007).

Once these initial divisions had been determined, additional resources were provided to boost the samples in various ways across each of the four countries: (1) an extra 35 disadvantaged wards were selected in England; (2) the sample in Wales was increased to 3,000 children and the additional 1,500 were to be drawn from disadvantaged wards only; (3) the Scottish sample was boosted by 1,000, to be split evenly between advantaged and disadvantaged wards; and (4) a further 500 children were allocated to Northern Ireland, all to be selected from disadvantaged wards. The boosted samples resulted in a total target sample of 20,646 children: 13,146 in England, 3,000 in Wales, 2,500 in Scotland, and 2,000 in Northern Ireland (Smith *et al.*, 2007).

### 3.3.4.4 Sample frame and recruitment

Once the final wards were chosen, a list of all 9-month-old infants residing in these areas was required (Smith *et al.*, 2007). The Department for Work and Pensions (DWP) provided access to the Child Benefit Register as a sampling frame, from which lists of all eligible participants were generated (Joshi & Fitzsimons, 2016). If a child was of the eligible age and living within a relevant ward, DWP officials contacted the Child Benefit recipient (typically the child's mother) by mail and asked them to opt out if they did not wish to participate in the MCS. For recruitment purposes, an opt-out method has been deemed more inclusive of marginal and low-literacy respondents when compared with an opt-in method and has been shown to yield greater response rates (Hansen, 2014). At present, opt-out consent is no longer permitted under GDPR legislation, whereby consent must always be provided via a clear affirmative act (e.g. ticked box, written/oral agreement) which establishes a freely given, specific, informed, and unambiguous signal of the participant's agreement to the processing of their personal data. Thus, silence, pre-ticked boxes or inactivity no longer constitute consent (GDPR.EU, 2018).

A total of 18,552 families were successfully recruited to the MCS, which, after allowing for 246 sets of twins and 10 sets of triplets, included 18,818 cohort children (Hansen, 2014). This final sample represents a response rate of 72% of all eligible families with children living in the sampled wards at age 9 months, and 81% of (assumed) eligible cases issued by the DWP for fieldwork purposes (Joshi & Fitzsimons, 2016).

### 3.3.4.5 Timing of additional waves

As can be seen in Table 3.6, there have been seven rounds of data collection thus far, with MCS follow-ups occurring approximately every 2–3 years (Joshi & Fitzsimons, 2016). Three online surveys regarding the COVID-19 pandemic were also administered to MCS participants between May 2020 and March 2021, and Wave 8 (age 22 years) is currently in development.

**Table 3.6** Timing of MCS data collection waves and response rates

| Data collection wave | Year | Age | Participants | Response rate |
|---|---|---|---|---|
| Wave 1 | 2001-2002 | 9 months | 18,818 | 96% |
| Wave 2 | 2003–2004 | 3 years | 15,808 | 81% |
| Wave 3 | 2006 | 5 years | 15,460 | 79% |
| Wave 4 | 2008 | 7 years | 14,043 | 72% |
| Wave 5 | 2012–2013 | 11 years | 13,469 | 69% |
| Wave 6 | 2015–2016 | 14 years | 11,872 | 61% |
| Wave 7 | 2018–2019 | 17 years | 10,757 | 56% |
| Wave 8 | 2022–2023 | 22 years | In development | N/A |

***Note:*** Response rates are given as a percentage of families participating in each wave out of the 19,244 families ever interviewed – no adjustment for death or emigration. Participants refer to cohort members (not families).

Source: CLS, 2021a; Joshi & Fitzsimons, 2016

### 3.3.4.6 Retention processes

The longitudinal pattern of response observed throughout the MCS is intricate, with attrition, re-entry, and a modest number of late entrants (n=702) at Wave 2 who were not included at Wave 1 because they had not yet been registered as living at an eligible address by the DWP (Connelly & Platt, 2014; Joshi & Fitzsimons, 2016). As described by Joshi and Fitzsimons (2016), "participant response in the MCS is not merely a one-way drain of permanent losses to follow-up; this is particularly evident between the second and third wave when 1,444 families returned" (p. 416). In fact, by Wave 5, 54% of the families who had ever participated had taken part in all five waves, and a further 20% had responded intermittently between waves (Joshi & Fitzsimons, 2016).

Naturally, a small number of families have permanently lost eligibility across waves owing to emigration or death of the child or parent. There has also been non-response at every round of data collection due to non-contact and withdrawal from the study. To counteract such non-response, the MCS has employed extensive office-based and field-based techniques to keep track of cohort families, such as the use of a freephone number, email address and website (Connelly & Platt, 2014). 'Keep in touch' leaflets – including updates, research findings and contact information self-reply cards – are circulated to cohort families once a year to help them feel more connected to the study (Wallace *et al.*, 2013). When families have moved but have failed to provide updated contact details, they are followed up through post office, electoral and telephone records, and via 'stable contacts' (e.g. friends, family members) whose details they have previously supplied. Regarding field-based procedures, interviewers would also contact the new property occupiers, neighbours, estate agents and other local sources in an attempt to locate the cohort member's new address (Connelly & Platt, 2014). However, it is important to note that such retrospective tracking would likely not be permissible nowadays under the GDPR. Moreover, the MCS has made sustained efforts to maintain positive relationships with study families, such as offering regular feedback and working on ways to alleviate respondent burden. As in the case of the GUI study, children have been given small gifts as tokens of appreciation, but also like GUI, there are no cash incentives for participation in the study (Joshi & Fitzsimons, 2016).

While the aforementioned methods have been critical in ensuring regular contact and engagement with cohort members outside of data collection waves, rates of non-response (both withdrawal and non-contact) for each wave have been consistently higher for families in ethnic minority or disadvantaged wards compared with those living in advantaged wards for each of the four UK countries (Connelly & Platt, 2014). For example, stratum-specific refusal rates for participants in ethnic or disadvantaged wards at Wave 4 were consistently higher when compared with those in advantaged wards across all UK countries; Northern Ireland had the highest refusal rate (23%), whereas those in advantaged wards in England had the lowest rate (14%) (Ketende, 2010). To address these differential patterns of response, weights were also provided in the accumulated data to account for inter-wave attrition (Connelly & Platt, 2014).

## 3.4 Conclusion

The commissioning of a longitudinal birth cohort study should prioritise practical feasibility of any proposed design and avoid design features that are overly ambitious or excessively burdensome to participants. Moreover, the effective design of a future birth cohort study needs to be guided by clear scientific and policy goals (Sullivan, Joshi & Williams, 2020). Bearing in mind the multitude of design considerations involved, this chapter began with a narrative synthesis of the academic literature regarding the main design features associated with the development of a birth cohort study. A review of international research practices emphasised the utmost importance of employing a predetermined and rigorous approach to sampling, alongside the need to actively ensure representativeness in order to avoid underestimating the prevalence of core outcomes among policy-relevant subgroups within the target population. Lessons for recruitment arising from stalled birth cohorts in the UK and the USA underscored the importance of balancing research ambition against burden on participating families. Novel design innovations (e.g. remote data collection, increased record linkage) currently being trialled in birth cohort studies in their early stages (e.g. GenV in Australia; the Early Life Cohort Feasibility Study (ELC-FS) in the UK) highlight this shift towards actively decreasing participant burden amid growing declines in response rates to population-based studies. Further design considerations pertaining to data protection and child welfare legislation were discussed. The literature review concluded with a brief evaluation of four potential sampling frames that might be implemented when designing a new birth cohort study within the Irish context.

The review was followed by an in-depth exploration of the specific design features chosen by the foundation case (GUI) and three additional case studies (the DNBC, ELFE and MCS). The analysis of the four case studies highlighted that response rates can fluctuate enormously across different waves and subpopulation groups. Design features, retention efforts and extraneous factors appear to affect willingness to participate at each wave. This is further evidenced by the varied response rates seen across case studies despite a similar frequency of data collection (e.g. GUI: 88% at age 3 years; MCS: 81% at age 3 years; ELFE: 67% at age 3.5 years). Regarding the main design considerations that could arise during the study development process, the following topics were covered for each case study: cohort age; sampling frame selection; sampling size calculation; ensuring ongoing representativeness; the timing of additional waves; and retention strategies. The rationale and experience of executing each of these chosen design features will help inform the development of a new birth cohort study in Ireland. See Chapter 6 for an in-depth synthesis of the concrete design feature options for the Irish context, alongside the associated benefits and challenges. The next chapter will explore the different data analysis techniques used in each of the four case studies, alongside a synthesis of core outputs (e.g. descriptive and analytic findings, key findings, technical documents, microdata) and mechanisms for maximising the data (e.g. conferences, workshops, access to microdata, documentation to support data utilisation).

# Appendices

## Appendix 3A

**Table 3.7** Summary of key features across the four named birth cohort studies

| | GUI | DNBC | ELFE | MCS |
|---|---|---|---|---|
| **Cohort age** | Birth cohort; 9 months | Pregnancy cohort; recruited between 6 and 12 weeks' gestation | From birth; infants recruited at maternity units | Birth cohort; 9 months |
| **Overall sample size** | 11,134 | 100,421 | 18,329 | 19,244 |
| **Sample size (at latest wave)** | 8,032<br>(Wave 5; 2018) | 25,898<br>(COVID-19 Survey; 2020) | 5,000<br>(COVID-19 Survey; 2020) | 10,757<br>(Wave 7; 2019) |
| **Sampling frame** | Child Benefit Register | Danish GPs | Maternity units in metropolitan France | Child Benefit Register |
| **Oversampling** | ● Non-marital births<br>● Non-national participants | ✗ | ✗ | ● Areas with high ethnic minority concentrations<br>● Socially disadvantaged areas |
| **Boosted samples** | ✗ | ✗ | ✗ | ✓ |
| **Use of weights** | ✓ | ✗ | ✓ | ✓ |
| **Timing of additional waves** | Five waves (every 2–3 years) | 12 waves (sporadically spread apart) | Multiple modes/sub-studies throughout childhood | Seven waves (every 2–3 years) |

# Chapter 4: Data analysis and outputs

## 4.1 Chapter overview

This chapter begins with an overview of a diverse range of findings from international longitudinal birth cohort studies. Following this review, the chapter focuses on the types of analyses produced by the four named birth cohort case studies: GUI, MCS, DNBC and ELFE. This section of the chapter is organised by analysis type: analyses produced by the study team are followed by external (non-study team) research. Issues of data access and strategies for maximising data use are discussed for each case study. The chapter concludes with an overview of different ways that findings from the birth cohort case studies have informed national policy and public debate.

## 4.2 Findings from birth cohort studies: An overview

Birth cohort studies such as GUI, Growing Up in Scotland, Growing Up in New Zealand, Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC), and the MCS are broad and multidisciplinary studies that focus on the whole life of the child from a social, developmental, economic, and health perspective. They have reported a range of findings across domains, starting from pregnancy, through the prenatal period, birth, infancy, and early childhood. Outlined below are some illustrative examples of the diverse range of findings generated, with data from different birth cohort studies, not just the four studies that this report focuses on. These findings are published either as analyses by the study teams themselves or those produced by external researchers.

Many studies have generated research findings using cohort study data concerning maternal health and the factors that impact the perinatal and early life health of children; for example, the MoBa study (Magnus *et al.*, 2006), the Growing Up in Scotland study (Bromley & Cunningham-Burley, 2010), and the EDEN study on the pre- and early postnatal determinants of child health and development (Heude *et al.*, 2016). Some studies have reported diverse findings regarding the impacts of maternal lifestyle on the early lives of children (e.g. smoking, alcohol consumption, physical activity, and medication use during pregnancy). These include the Slovak

PCB Study (Slovakia), Born in Bradford (UK), and the INUENDO cohort (Sweden, Poland, Greenland, and Ukraine). Regarding parenting and nutrition, researchers using Growing Up in Scotland data found that the majority (88%) of parents with children aged 10 years who were overweight perceived their child's weight as normal (Bradshaw & Hinchliffe, 2018). These findings, among others, resulted in recommendations to engage with and educate parents about early childhood overweight and obesity to improve and maintain childhood health and weight.

Several birth cohort studies have reported findings related to maternal obstetrics, including fertility treatment, number of times giving birth, waiting time to pregnancy, mode of delivery, and prenatal diagnostics. These include ALSPAC (UK) and the Aarhus Birth Cohort Biobank (Denmark). For instance, women who underwent fertility treatment and who also had a diagnosis of endometriosis were at an increased risk of preeclampsia, early birth, and caesarean section, irrespective of the type of fertility treatment received (Glavind *et al.*, 2017).

Analyses from other studies such as the Trieste Cohort (Italy) have provided a range of findings on the impact of maternal environmental exposures on child health (e.g. occupational hazards, outdoor air pollution, indoor air pollution, and environmental exposures and their impact on allergies, skin conditions such as eczema and psoriasis, and respiratory health). For example, Neuman *et al.* (2012) performed a pooled analysis of eight European birth cohort studies and found that, among preschool-aged children, there was a significant risk for developing a respiratory illness due to maternal smoking during the first trimester of pregnancy.

With regard to socioeconomic and living conditions, Growing Up in Scotland study researchers found that families studied in the period 2011–2012 had a lower household income than families studied in the period 2005–2006. Moreover, they found that more than one-half of the families with a baby in 2011–2012 had no security in the form of savings or investments (Bradshaw *et al.*, 2013). The Growing Up in New Zealand study team reported a range of findings related to residential mobility status. For example, it was found that moving home is a frequent occurrence for families, particularly in the first 1,000 days of a baby's life, due to short or insecure rental contracts, parental employment status, and neighbourhood features (Morton *et al.*, 2014). Regarding service provision, researchers using Growing Up in Australia:

(LSAC) data found that a large proportion of children with psychosocial adjustment issues such as hyperactivity were not in receipt of appropriate services to support their needs (e.g. psychiatric and behavioural therapy services and psychologists) (Warren, Quinn & Daraganova, 2020). As is evident from the findings outlined above, birth cohort studies report on an abundance of salient outcomes across key domains, from pregnancy into early childhood.

## 4.3 Analyses by cohort study teams

In this section we report on the types of analyses produced by cohort study teams. These tend to focus on the most recent wave of the study at the time these analyses were carried out, but may also include longitudinal analyses of multiple waves. Many of the examples below focus on the very first wave of the studies.

### 4.3.1 Growing Up in Ireland study

The GUI study team produces four main types of reports:

- **Key findings reports:** short reports for each wave of data collection highlighting key findings across the main domains. These are often published swiftly after each wave. For example:
    - Growing Up in Ireland Study Team (2011) *Key findings: Infant Cohort (at 9 months)* (Pregnancy and Birth No.1). Dublin: Economic and Social Research Institute, Trinity College Dublin, and Office of the Minister for Children and Youth Affairs.
    - Growing Up in Ireland Study Team (2011) *Key findings: Infant Cohort (at 9 months)*. (Infant Health No. 2.). Dublin: Economic and Social Research Institute, Trinity College Dublin, and Office of the Minister for Children and Youth Affairs.
    - Growing Up in Ireland Study Team (2011) *Key findings: Infant Cohort (at 9 months)*. (Childcare and Parenting Support. No. 3.) Dublin: Economic and Social Research Institute, Trinity College Dublin, and Office of the Minister for Children and Youth Affairs.
    - Murray, A., McNamara, E., O’Mahony, D., Smyth, E. and Watson, D. (2021) *Growing Up in Ireland: Key findings from the special COVID-19 survey of Cohorts ’98 and ’08*. Economic and Social Research Institute.
- **Descriptive reports:** provide more detailed analysis of data from each wave and from several waves longitudinally. These reports showcase the data, its policy relevance, and the potential for further analysis. For example:
    - Williams, J., Murray, A., McCrory, C. and McNally, S. (2013) *Growing Up in Ireland national longitudinal study of children: Development from*

  *birth to three years infant cohort.* Report 5. Dublin: Department of Children and Youth Affairs.
  - Murray, A., McNamara, E., Thornton, M., Williams, J. and Smyth, E. (2019) *Growing Up in Ireland – The lives of 5-year-olds.* Dublin: Department of Children and Youth Affairs, Economic and Social Research Institute, and Trinity College Dublin.
  - McNamara, E., Murray, A., O'Mahony, D., O'Reilly, C., Smyth, E. and Watson, D. (2021) *Growing Up in Ireland: The lives of 9-year-olds of cohort '08*. Report No. 10. Dublin: Department of Children and Youth Affairs, Economic and Social Research Institute, and Trinity College Dublin.

- **Thematic reports:** focus on detailed analyses of specific, policy-relevant topics. For example:
  - McGinnity, F., Murray, A. and McNally, S. (2013) Growing Up in Ireland: National Longitudinal Study of Children: Mothers' return to work and childcare choices for infants in Ireland. Infant Cohort Research Report No. 2. Dublin: Department of Children and Youth Affairs.
  - Layte, R. and McCrory, C. (2014) Growing Up in Ireland: Maternal health behaviours and child growth in infancy. Infant Cohort Research Report No. 4. Dublin: The Stationery Office.
  - Nolan, A. and Layte, R. (2017) Understanding use of general practitioner services among children in Ireland. Dublin: The Stationery Office.

- **Technical reports:** Primary resources for researchers planning to use GUI study data (e.g. pilot reports, design reports, summary guides to each wave of data, and questionnaires). All reports are available on the GUI website, including, for example, those pertaining to Wave 1 of GUI Cohort '08 (see Table 4.1). Section 4.3.1.1 offers examples of the diversity of findings coming out of Wave 1 from primary research published by the GUI research team.

**Table 4.1** GUI research reports related to GUI Cohort '08 at 9 months[20]

| |
|---|
| Key findings reports – Wave 1 at 9 months:<br>• Pregnancy and birth<br>• Infant health<br>• Childcare and parenting support |
| Descriptive, thematic, and technical reports – Wave 1 at 9 months:<br>• *Maternal Health Behaviours and Child Growth in Infancy* (thematic)<br>• *Report on the Qualitative Study of Infants and their Parents at Wave 1* (technical/ descriptive)<br>• *Parenting and Infant Development* (thematic)<br>• *Mothers' Return to Work and Childcare Choices for Infants in Ireland* (thematic)<br>• *The Infants and their Families* (descriptive) |

Source: Growing up in Ireland, 2022

---

[20] Official Publications from GUI Cohort '08 (Infant Cohort) can be found at https://www.growingup.ie/growing-up-in-ireland-publications-on-the-infant-cohort/.

#### 4.3.1.1 Examples of findings from GUI study reports

A thematic GUI report by Nolan and Layte (2017) using data from the first two waves of GUI Cohort ’08 (i.e. when the children were aged 9 months and 3 years, respectively), examined the use of GP services among children in Ireland. The report found that children from lower-income households who did not have a full Medical Card or GP Visit Card had a lower number of visits (mainly at age 9 months) when compared with children from higher-income homes. This finding identified the possibility that economic limitations reduce lower-income families’ ability to access GP services for their children. A descriptive GUI report by Murray *et al.* (2019) utilised data from Wave 1 of the GUI Cohort ’08 and focused on the lives of 5-year-olds across a wide range of domains and topics. Among other findings, Murray and colleagues observed links between children’s technological device use and risk of obesity. They reported that higher rates of screen time (including television, smartphones and other devices) were associated with higher BMI status and greater risk of overweight or obesity among GUI Cohort ’08 children at age 5 years. In addition, children who engaged in 3 or more hours of screen time per day were far more likely to consume sweets and snacks than children who had less screen time. This finding did not vary by household income.

Studying GUI cohort members during the COVID-19 pandemic has also provided researchers with a wealth of essential data on the well-being and experiences of children at a time of worldwide crisis. For instance, the GUI COVID-19 survey has offered important insights into children’s education during the pandemic. Approximately two-thirds of GUI Cohort ’08 children (aged 12 years), who started their first year of secondary education in September 2020 after the initial national lockdown, reported difficulties related to school life (e.g. challenges settling in, difficulties with schoolwork and homework) (Murray *et al.*, 2021). More generally, the COVID-19 survey focused on participants’ experiences of the pandemic, including the effects on learning and employment, changes in free-time activities, personal experience of COVID-19, sources of information about COVID-19, and emotional well-being.

### 4.3.2 Millennium Cohort Study

The MCS study team provides detailed user guides, questionnaires, and other survey documentation, including technical reports on data collection procedures. These are publicly available via the Centre for Longitudinal Studies (CLS) website[21]. The MCS study team and other members of the CLS research team publish their own analyses of the MCS data in the CLS working paper series[22] and in academic journals. They also publish summaries of their MCS-based analyses in CLS's briefing papers series.[23] These tend to focus on a specific topic, similar to the GUI's key findings or thematic reports. For example, a recent briefing paper by Patalay and Fitzsimmons (2021) analysed the prevalence of and inequalities in psychological distress, self-harm and attempted suicide among 17-year-old MCS cohort members. However, unlike the GUI team, the MCS study team regularly publishes detailed descriptive statistics from every sweep of the study. One exception is the report published by the MCS research team related to Sweep 1 or the first survey of the MCS (Dex & Joshi, 2004). Findings within this report relate to household structure and characteristics, ethnic identity, partnerships and parenthood, wider family, pregnancy, delivery and labour, babies' health and development, parenting and parents' psychosocial adjustment, parents' health, parental employment and education, childcare, income and benefits, housing and the area, and citizenship.

### 4.3.3 Danish National Birth Cohort

There are no research reports or key findings reports directly available for the DNBC study on its study website. Offered instead is the study bibliography (addressed below in Section 4.4.4) and a selection of DNBC publications on the cohort background and methods (outlined in Section 4.5.3.1 DNBC technical documents)

### 4.3.4 Étude Longitudinale Française depuis l'Enfance (ELFE)

The ELFE study does not provide key findings reports from the study data. While the GUI study offers a clear overview of the different outputs produced by the GUI study team and lists additional reports generated by external researchers using GUI data,

---

[21] https://cls.ucl.ac.uk/publications-and-resources/

[22] https://cls.ucl.ac.uk/working_papers/

[23] https://cls.ucl.ac.uk/briefings_impact/

ELFE does not provide guidance about report types. Table 4.2 lists the 15 reports and working documents available on the study website. The titles of the papers have been translated for the purpose of this report.[24] The reports are divided into reports for French Government Departments, including six reports at the request of the General Directorate of Health, three requested by the French National Family Allowance Fund (Caisse Nationale d'Allocations Familiales; CNAF), two requested by Public Health France/the Environmental Health Department, and one report requested by the French Public Health Research Institute.

**Table 4.2** ELFE reports and working documents

- Gassama, M., Bernard, J., Dargent-Molina, P. and Charles, M.-A. (2018).*Physical activities and use of screens at the age of 2 years among children of the ELFE cohort*. Statistical analysis and report prepared on request and with the financial support of the Direction Générale de la Santé.
- Samuel, O., Brachet, S., Brugeilles, C., Pélage, A., Paillet, A. and Rollet, C. (2014). *Production and parental reception of gender norms: mothers and fathers facing their baby girls and boys*. Report for the CNAF ("Parentality(ies): production and reception of norms" call for projects).
- Berton, F., de Bony, J., Bureau, M.-C., Jung, C., Rist, B. and Touahria-Gaillard, A. (2015). *Being a parent in the face of institutions: norms of parenthood and paradoxical injunctions in public action*. CNRS/CNAM UMR3320 final report for the CNAF.
- Moguerou, L., Eremenko, T., Thierry, X. and Prigent, R. (2016). *Profiles, paths and experiences of immigrant single-parent families*. Final report for the CNAF.
- Public Health France/Environmental Health Department: Dereumeaux, C., Guldner, L., Saoudi, A., Pecheux, M., de Crouy-Chanel, P., Bérat, B., Wagner, V. and Goria, S. (2016). *Impregnation of women pregnant by environmental pollutants in France in 2011. Perinatal component of the national biomonitoring program implemented within the ELFE cohort*. Volume 1: Organic Pollutants.
- Public Health France/Environmental Health Department: Dereumeaux, C., Fillol, C., Saoudi, A., Pecheux, M., de Crouy-Chanel, P., Bérat, B, Wagner, V. and Goria, S. (2017) *Impregnanation of pregnant women by environmental pollutants in France in 2011. Perinatal component of the national biomonitoring program implemented within the ELFE cohort*. Volume 2: Metals and Metalloids.
- De Lauzon, B. (2018). *Food for children from birth to 1 year old. Synthesis of the results of the ANR SOFI project*. General Directorate of Health.
- Gassama, M. and Charles, M. A. (2018). *Food for children aged 1 to 2 years*. ELFE Cohort.
- Pailhé, A., Solaz, A. and Tô, M. (2018). *Can daddies learn how to change nappies? Evidence from a Short Paternity Leave Policy*. Collection of INED working documents, No. 240.

---

[24] For the original French titles of the reports, see https://www.elfe-france.fr/en/the-research/publications/reports-and-working-documents/

- Gassama, M., Bernard, J., Dargent-Molina, P. and Charles, M. A. (2018). *Physical activities and use of screens at the age of 2 and 3.5 years among children of the Elf cohort*. Report for the General Directorate of Health.
- Gassama, M., Heude, B., Forhan, A. and Charles, M. A. (2018). *Growth of children in the ELFE cohort from 0 to 2 years old*. Report for the General Directorate of Health.
- Lepeule, J. (2018*). PATer Final Report – N°EST-2013-216, PATer, Air pollution on French territory: Modeling and health effects*. Project funded by ANSES as part of the PNR EST.
- Milcent, K., Bois, C. and Charles, M. A. (2019). *Health check-up of children aged 3-4 years: Impact of the first primary prevention examination in kindergarten*. Report for the call for research projects of the Public Health Research Institute.
- Gassama, M., Milcent, K., Bois, C., Dufourg, M. N. and Charles, M. A. (2019) *Impact of the first primary prevention examination in kindergarten - Partnership between the departmental services of PMI and the ELFE cohort*. Report to the request and with the support of the General Directorate of Health.
- Gassama, M., Milcent, K., Dufourg, M. N. and Charles, M. A. (2019). *Hearing screening in 2011 and prevalence of hearing disorders in the ELFE cohort*. Report for the General Directorate of Health.

## 4.3.5 Generation Victoria

As the Generation Victoria (GenV) study has just started data collection, there are no available analyses of the study results to date. However, there are several journal articles produced by the study team in collaboration with other researchers which outline some key features of the study design (see Table 4.3). There are also several working papers[25] written by the GenV study team which focus on the methodological aspects of the project.

**Table 4.3** GenV study background and design papers

- Sung, V., Williams, K., Perlow, E., Hu, Y. J., Ahern, S., Said, J. M., Karanatsios, B., Hopper, J. L., McNeil, J. J., Donnan, L., Goldfeld, S. and Wake, M. (2021) 'Enhancing Value and Uptake for Whole-Population Cohorts of Children and Parents: Methods to Integrate Registries into the Generation Victoria Cohort', *Children*, 8(4), p. 285. https://doi.org/10.3390/children8040285
- Wake, M., Hu, Y. J., Warren, H., Danchin, M., Fahey, M., Orsini, F., Pacilli, M., Perrett, K. P., Saffery, R., & Davidson, A. (2020) 'Integrating trials into a whole-population cohort of children and parents: statement of intent (trials) for the Generation Victoria (GenV) cohort', *BMC Medical Research Methodology*, 20(1), pp. 1–15. https://doi.org/10.1186/s12874-020-01111-x
- Wang, J., Hu, Y. J., Clifford, S., Goldfeld, S. and Wake, M. (2021) 'Selecting life course frameworks to guide and communicate large new cohort studies: Generation Victoria (GenV) case study', *Journal of Developmental Origins of Health and Disease*,*12*(6),c pp. 1–20. https://doi.org/10.1017/S2040174420001245

---

[25] https://www.genv.org.au/for-researchers/working-papers/

Several useful lessons have arisen from the above GenV papers. First, Wang *et al.* (2021) demonstrated the importance of a well-considered strategy and transparent selection process when selecting a conceptual framework to underpin and grow with such a large-scale birth cohort study. Wang and colleagues (2021) also illustrated the value of using visual aids to communicate complex points to non-expert audiences in a user-friendly way. Second, Wake *et al.* (2020) recommend utilising the huge breadth of large-scale birth cohort studies to test and explore the feasibility and limits of integrating trials focused on children and young adults into birth cohorts. Finally, Sung *et al.* (2021) propose that GenV should be used to identify and develop new principles, methods, and guidance structures to integrate health registries, biosamples, phenotypes and self-reported participant measures into the one birth cohort study.

## 4.4 Analyses led by external (non-study team) researchers

In this section we provide an overview of the types of independent external research conducted primarily by academics using data from the GUI study and the named birth cohort studies. This research tends to exploit the longitudinal nature of the studies and, as such, uses more advanced data analysis techniques (e.g. panel regression, growth curve modelling).

### 4.4.1 Search strategy for secondary analyses

A two-step process was undertaken to identify published empirical research papers which utilised GUI, MCS, DNBC or ELFE data. First, a rapid evidence assessment (Garritty *et al.*, 2021) exercise was undertaken by the review team of the top 10 journals in sociology, social science (public health), health and medicine, epidemiology, psychology, and economics. The review team then searched these journals by article title and abstract for keywords associated with the naming conventions of the case studies (i.e. Growing Up in Ireland, GUI, Millennium Cohort Study UK, MCS, Danish National Birth Cohort, DNBC, Étude Longitudinale Française depuis l’Enfance, ELFE).

Next, highly cited empirical research papers per case study were identified through the Web of Science (WoS) core collection database, again using a retrieval query of the keywords associated with the naming conventions of the case studies and defining the document type as ‘article and review’. A minimum of 20 papers were identified per case study. The following bibliometric indicators were recorded for a minimum of the top 20 most highly cited research publications per case study: authors, article title, source title, abstract, primary author contact, number of times cited (all databases), publication year, volume, issue, and DOI. Examples of highly cited papers per case study are identified in Sections 4.4.2-4.4.6.

### 4.4.2 Growing Up in Ireland

The GUI website includes a repository of all GUI study publications, and also maintains a database of research conducted by external researchers using GUI data. The bibliography of publications by external researchers[26] on the GUI website identified 224 secondary research papers using GUI data for their study (accessed April 2022). A search of titles and abstracts via WoS, returned 167 results (see Appendix 4B for WoS search results for all named case studies). Cross-checking these entries is outside the scope of this report, so there may be an overlap between the WoS results and those on the GUI website.

Research associated with the GUI data is most frequently conducted in the areas of social science, health, epidemiology, education, child development, and psychiatry. For example, a recent secondary data analysis paper published by Mohan (2021) used multi-wave data from both cohorts of the GUI study to analyse the public health risk of energy poverty for children in Ireland. The author observed higher rates of respiratory illness and childhood wheezing in children who experienced household energy poverty. This work identified the policy implications of this finding, including strategies to focus on health and social inequalities and the promotion of energy justice. Another study by Jabakhanji *et al.* (2017) used Wave 1 and Wave 2 data from GUI Cohort ’08 to assess the link between social class and economic recession on obesity rates among 3-year-old children living in Ireland. The researchers found that increased rates of obesity in this group were significantly associated with families’

---

[26] https://www.growingup.gov.ie/information-for-researchers/all-publications-using-growing-up-in-ireland-data/

perceived negative impact of the recession, but not with social class. Briody (2021) used data from the first three waves of the GUI study to analyse the effects of parental unemployment on childhood obesity. McCrory *et al.* (2019) used multi-wave data from both GUI cohorts to study trajectories in BMI in boys and girls by maternal educational attainment.

To offer further insight into where GUI secondary research is published, the range of journals with highly cited research articles using GUI data are listed in Table 4.4. Table 4.5 provides examples of these highly cited research articles using GUI data.

**Table 4.4** The range of journals with highly cited research articles using GUI data

| | |
|---|---|
| • *BMC Pediatrics* | • *Journal of Paediatrics and Child Health* |
| • *British Journal of General Practice* | • *Journal of Physical Activity and Health* |
| • *Child: Care, Health and Development* | • *The Lancet Psychiatry* |
| • *Child Language Teaching and Therapy* | • *Maternal and Child Health Journal* |
| • *Economics & Human Biology* | • *Paediatric and Perinatal Epidemiology* |
| • *Epidemiology and Psychiatric Sciences* | • *Pediatrics* |
| • *European Journal of Public Health* | • *Pediatric Research* |
| • *European Journal of Special Needs Education* | • *Personality and Individual Differences* |
| • *International Journal of Obesity* | • *PLOS ONE* |
| • *Irish Educational Studies* | • *Public Health Nutrition* |
| • *Journal of Abnormal Child Psychology* | • *Research in Developmental Disabilities* |
| • *Journal of Epidemiology & Community Health* | • *Social Science & Medicine* |
| • *Journal of Biosocial Science* | • *Social Psychiatry and Psychiatric Epidemiology* |

**Table 4.5** Examples of highly cited research articles using GUI data

| |
|---|
| Keane, E., Layte, R., Harrington, J., Kearney, P. M. and Perry, I. J. (2012) 'Measured Parental Weight Status and Familial Socio-Economic Status Correlates with Childhood Overweight and Obesity at Age 9', *PLOS ONE*, 7(8), p. e43503. https://doi.org/10.1371/journal.pone.0043503 (cited 93 times) |
| McCrory, C. and Layte, R. (2012) 'Breastfeeding and risk of overweight and obesity at nine-years of age', *Social Science & Medicine*, 75(2), pp. 323–330. https://doi.org/10.1016/j.socscimed.2012.02.048 (cited 66 times) |
| McCrory, C. and McNally, S. (2013) 'The effect of pregnancy intention on maternal prenatal behaviours and parent and child health: results of an Irish cohort study', *Paediatric and Perinatal Epidemiology*, 27(2), pp. 208–215. https://doi.org/10.1111/ppe.12027 (cited 57 times) |

### 4.4.3 Millennium Cohort Study

MCS data are used primarily by researchers working in the disciplines of health, education, social studies, and economics to investigate how the life course is impacted by a range of circumstances and conditions (Connelly & Platt, 2014; Smith & Joshi, 2002). Since 2000, findings associated with MSC study data have helped identify how health and development are impacted by a complex range of factors in early life. The UK CLS's website identified 1,257 research papers associated with MCS data[27]. A search of titles and abstracts conducted via WoS returned 358 results. Examples of the breadth of findings produced by research teams and other data users are outlined below. These findings have been selected to showcase the diversity of findings across varied topics that the MCS has produced.

Quigley, Kelly and Sacker (2007) used MCS data to measure the impact of breastfeeding on emergency medical care for diarrhoea and lower respiratory tract infections between birth and age 8 months in the UK. This study found that exclusively breastfed babies were less likely to be hospitalised for both conditions. Indeed, more than 50% of hospitalisations related to diarrhoea and nearly 30% of hospitalisations for respiratory tract infections could have been prevented by exclusive breastfeeding. Although having a marginally lower impact, partial breastfeeding was also found to have clear health benefits for babies in comparison to those who were not breastfed. The researchers also observed that prolonged

---

[27] https://www.bibliography.cls.ucl.ac.uk/

length of breastfeeding positively impacted babies' health. Fitzsimons *et al.* (2017) published results related to poverty dynamics and parental mental health and their impact on the mental health of children. This study showed that children aged 5 years and 11 years who experienced consistent levels of poverty or who transitioned into worsening levels of poverty experienced a higher rate of mental health problems, including hyperactivity, emotional issues, behaviour issues and peer-related problems. There was also a strong correlation between maternal mental health problems (more so than paternal mental health) and negative impacts on children's mental health.

Yang, Petersen and Qualter (2020) observed that participants aged 14 years reported rates of loneliness that are often associated with older age groups, and that those affected by feelings of loneliness often struggle in silence as they keep these feelings hidden. In addition, this study reported that loneliness was not singularly related to a lack of friends but was also influenced by engaging with people and groups who adolescents perceived to be harmful to their well-being. Girls were more likely to experience feelings of loneliness than boys, as were participants from white or mixed ethnic backgrounds. Moreover, participants living in Wales reported greater feelings of loneliness compared to those living in England, Northern Ireland, and Scotland.

Cited secondary research associated with MCS data was found in a range of journals, examples of which are listed in Table 4.6 and Table 4.7.

**Table 4.6** The range of journals with highly cited research published using MCS data

| | |
|---|---|
| • *American Journal of Epidemiology* | • *International Journal of Epidemiology* |
| • *BMC Public Health* | • *International Journal of Obesity* |
| • *European Journal of Developmental Psychology* | • *International Journal of Social Research Methodology* |
| • *European Journal of Epidemiology* | • *Public Health Nutrition* |
| • *Ethnic and Racial Studies* | • *Oxford Review of Education* |
| • *International Journal of Behavioral Development* | • *Journal of Official Statistics* |

**Table 4.7** Examples of highly cited papers related to MCS data

| |
|---|
| Kiernan, K. E. and Huerta, M. C. (2008) 'Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood', *The British Journal of Sociology*, *59*(4), pp. 783–806. https://doi.org/10.1111/j.1468-4446.2008.00219.x (cited 258 times) |
| Griffiths, L. J., Cortina-Borja, M., Sera, F., Pouliou, T., Geraci, M., Rich, C., Cole, T. J., Law, C., Joshi, H., Ness, A. R., Jebb, S. A., & Dezateux, C. (2013) 'How active are our children? Findings from the Millennium Cohort Study', *BMJ Open*, 3(8), p. e002893. http://dx.doi.org/10.1136/bmjopen-2013-002893 (cited 233 times) |
| Rich, C., Geraci, M., Griffiths, L., Sera, F., Dezateux, C., & Cortina-Borja, M. (2013b) 'Quality Control Methods in Accelerometer Data Processing: Defining Minimum Wear Time', *PLOS ONE*, 8(6), p. e67206. https://doi.org/10.1371/journal.pone.0067206 (cited 161 times) |

## 4.4.4 Danish National Birth Cohort

The core research outputs associated with DNBC data focus on foetal and perinatal outcomes which may be affected by maternal health, well-being, and lifestyle during pregnancy (Olsen & Meder, 2014).

The DNBC study is primarily focused on understanding the causal link between exposures in perinatal/early life and disease development across the life span; thus, research is most frequently conducted in the disciplines of medicine, epidemiology, public health, environmental health, and psychiatry. As of February 2022, the DNBC's official bibliography, which is hosted on the National Library of Medicine website,[28] identified 688 secondary research papers associated with DNBC data. A search of titles and abstracts via WoS returned 257 results for the DNBC study.

Examples of the breadth of health-related findings associated with the DNBC study are outlined below. Since its launch, DNBC research findings have focused on domains of lifestyle and substance use during pregnancy (e.g. alcohol, coffee, smoking, and nicotine); nutrition, diet, and physical activity during pregnancy; pregnant women's health; environmental toxins; infections during pregnancy; medication intake during pregnancy; and maternal mental health.

For example, Zhu *et al.* (2014) undertook a study to examine the effects of parental smoking during pregnancy utilising the DNBC data. This study found that both

---

[28] The DNBC's official bibliography can be found on the National Library of Medicine website: https://www.ncbi.nlm.nih.gov/.

maternal and parental smoking was linked to a higher risk of attention deficit hyperactivity disorder (ADHD) in children. However, maternal nicotine intake (i.e. smoking or nicotine replacements) had a stronger impact than paternal smoking. Zhu *et al.* (2004) used DNBC data to measure whether shift work and work-related stress were associated with late-term foetal loss. This study found that dedicated night shift work during pregnancy was correlated to late-term foetal loss in comparison with other types of shift work. Work-related stress was not found to be related to foetal loss. Liew *et al.* (2016) reported that maternal paracetamol intake during pregnancy was correlated with increased risk of autism spectrum disorder (ASD) with hyperkinetic symptoms, and that use over 20 weeks or more doubled the risk of ASD with hyperkinetic symptoms. These findings were based on DNBC data related to 64,322 children and their mothers over nearly 13 years.

Highly cited secondary research associated with DNBC data was found in a range of journals, examples of which are listed in Table 4.8 and 4.9.

**Table 4.8** The range of journals from highly cited research published using DNBC data

| | |
|---|---|
| • *Autism Research* | • *International Journal of Gynecology & Obstetrics* |
| • *Diabetes Care* | • *International Journal of Pediatric Obesity* |
| • *Environmental Health* | • *Journal of Allergy and Clinical Immunology* |
| • *Epidemiology* | • *Journal of Clinical Endocrinology & Metabolism* |
| • *European Child & Adolescent Psychiatry* | • *Journal of Child Psychology and Psychiatry* |
| • *European Journal of Epidemiology* | • *Journal of Occupational and Environmental Medicine* |
| • *Genetic Testing and Molecular Biomarkers* | • *Paediatric and Perinatal Epidemiology* |
| • *Human Reproduction* | • *PLOS ONE* |
| • *International Journal of Environmental Research and Public Health* | • *Public Health Nutrition* |
| • *International Journal of Epidemiology* | • *Scandinavian Journal of Work Environment & Health* |
| • *International Journal of Obesity* | • *Thyroid* |

**Table 4.9** Examples of highly cited papers related to DNBC data

| |
|---|
| Ajslev, T. A., Andersen, C. S., Gamborg, M., Sørensen, T. I. and Jess, T. (2011) 'Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics', *International Journal of Obesity*, 35(4), pp. 522–529. https://doi.org/10.1038/ijo.2011.27 (cited 576 times). |
| Zhu, Y., Olsen, S. F., Mendola, P., Halldorsson, T. I., Rawal, S., Hinkle, S. N., Yeung, E. H., Chavarro, J. E., Grunnet, L. G., Granström, C., Bjerregaard, A. A., Hu, F. B. and Zhang, C. (2017) 'Maternal consumption of artificially sweetened beverages during pregnancy, and offspring growth through 7 years of age: a prospective cohort study', *International Journal of Epidemiology*, 46(5), pp. 1499–1508. https://doi.org/10.1093/ije/dyx095 (cited 64 times). |
| Schmidt Morgen, C., Rokholm, B., Sjöberg Brixval, C., Schou Andersen, C., Geisler Andersen, L., Rasmussen, M., Nybo Andersen, A. M., Due, P. and Sørensen, T. I. (2013) 'Trends in prevalence of overweight and obesity in Danish infants, children and adolescents--are we still on a plateau?', *PLOS ONE*, 8(7), p. e69860. https://doi.org/10.1371/journal.pone.0069860 (cited 59 times). |

### 4.4.5 Étude Longitudinale Française depuis l'Enfance

The goal of the ELFE study is to examine the causes of child health, growth, development, and social skills from birth to adulthood using a multidisciplinary approach (Charles *et al.*, 2020; Vandentorren *et al.*, 2009). The ELFE study has published a range of results and findings across the domains of social science, health, and the environment. As of February 2022, the ELFE study's official bibliography identified 192 secondary research papers associated with the cohort data.[29] A search of titles and abstracts via WoS returned 82 results for the ELFE study. Research outputs associated with ELFE data are mostly published in the disciplines of epidemiology, public health, environmental health, and social sciences.

In the area of social science, the ELFE study has identified findings related to children's lived experience during the COVID-19 restrictions in France (Monnier *et al.*, 2021), low birthweight and its relationship with socioeconomic attributes (Grobon, Panico & Solaz, 2019), the effect of socioeconomic disparities on children's development (Panico, Tô & Thévenon, 2015), and the role of grandparents during early life (Thalineau & Nowik, 2018). Health research outcomes related to the ELFE study cohort include the sleep habits of 1-year-olds (Messayke *et al.*, 2020), the Bacillus Calmette–Guérin (BCG) vaccination before age 3 months (Guthmann *et al.*,

---

[29] The ELFE study's official bibliography can be found at https://www.elfe-france.fr/en/the-research/publications/academic-journals/.

2016), access to mental health services for mothers during pregnancy (Bales *et al.*, 2015), and the link between migrant background and substance use during pregnancy (Melchior *et al.*, 2015). Environmental research outcomes related to the ELFE cohort include exposure to microorganisms in the home (Rocchi *et al.*, 2015) and dietary exposures to pollutants (de Gavelle *et al.*, 2016).

ELFE study data have been used to support the case for a reduction in technological and mobile device use in young children. One study using the ELFE data reported that high rates of screen time among children aged 2–3 years were correlated with poor sleep patterns, learning disorders, and behavioural issues (Berthomier & Octobre, 2019). Another French study utilised ELFE cohort data to identify associations between environmental toxins and pollutants and maternal health during pregnancy (Cognez *et al.*, 2019). Results were statistically significant regarding the association between the use of pesticides and insect repellents during pregnancy and the risk of underdeveloped and malformed babies at birth. This study argues the case for limiting the use of such chemicals in general and for protecting women from these toxins during pregnancy. Highly cited secondary research associated with ELFE data were found in a range of journals, examples of which are listed in Table 4.10 and 4.11.

**Table 4.10** The range of journals with highly cited research using ELFE data

| | |
|---|---|
| • *Addictive Behaviors* | • *International Journal of Hygiene and Environmental Health* |
| • *BMC Pediatrics* | • *International Journal of Public Health* |
| • *Environment International* | • *Maternal & Child Nutrition* |
| • *Environmental Research* | • *Midwifery* |
| • *European Psychiatry* | • *Pediatric Allergy and Immunology* |
| • *International Journal of Epidemiology* | • *PLOS ONE* |

**Table 4.11** Examples of highly cited papers related to ELFE data

| |
|---|
| Dereumeaux, C., Saoudi, A., Pecheux, M., Berat, B., de Crouy-Chanel, P., Zaros, C., Brunel, S., Delamaire, C., le Tertre, A., Lefranc, A., Vandentorren, S. and Guldner, L. (2016) ‘Biomarkers of exposure to environmental contaminants in French pregnant women from the ELFE cohort in 2011’, *Environment International*, *97*, pp. 56–67. https://doi.org/10.1016/j.envint.2016.10.013 (cited 85 times). |
| Bales, M., Pambrun, E., Melchior, M., Glangeaud-Freudenthal, N. M., Charles, M. A., Verdoux, H. and Sutter-Dallay, A. L. (2015) ‘Prenatal psychological distress and access to mental healthcare in the ELFE cohort’, *European Psychiatry: The Journal of the Association of European Psychiatrists*, 30(2), pp. 322–328. https://doi.org/10.1016/j.eurpsy.2014.11.004 (cited 50 times). |
| Béranger, R., Hardy, E. M., Dexet, C., Guldner, L., Zaros, C., Nougadère, A., Metten, M. A., Chevrier, C., & Appenzeller, B. M. R. (2018) ‘Multiple pesticide analysis in hair samples of pregnant French women: results from the ELFE national birth cohort’, *Environment International*, *120*, pp. 43–53. https://doi.org/10.1016/j.envint.2018.07.023 (cited 47 times). |

### 4.4.6 Generation Victoria

There are currently no secondary analyses of GenV data, given that the GenV study has not yet completed its first wave of data collection.

## 4.5 Data access

The following section details the GUI and named birth cohort studies’ technical documents, data access arrangements, and methods for maximising data access and analysis.

### 4.5.1 GUI data access

#### 4.5.1.1 GUI technical documents

Alongside research reports detailing analyses of GUI data by the study team, a full collection of supporting data documentation is also available for each data wave. These consist of data summary guides, questionnaires, design reports, pilot reports, data codebooks, data dictionaries, variable naming guides, sample design and response rate overview documentation, and a derived variables guide for Researcher Microdata Files. For example, Table 4.12 lists technical documents that are available for the GUI Cohort ’08 at Wave 1 (age 9 months).

**Table 4.12** Technical documents for GUI Cohort '08 at age 9 months (GUI, 2022) [30]

- *Cohort '08 at 9 Months Codebook for Wave 1 of the Infant Cohort*
- *Data Dictionary for Wave 1 of the Infant Cohort (at 9 months)*
- *Questionnaires for Wave 1 of the Infant Cohort (at 9 months)*
- *Sample design and response in Wave 1 of the Infant Cohort (at 9 months)*
- *Summary Data Dictionary for Wave 1 of the Infant Cohort (at 9 months)*
- *Summary Guide to Wave 1 of the Infant Cohort (at 9 months)*
- *Variable Naming Conventions and Longitudinal Data Dictionary for Wave 1 and Wave 2 of Cohort '08*
- *Technical Report on the Qualitative Data from the Infant Cohort at Nine Months of Age*
- *Design, Instrumentation and Procedures for the Infant Cohort at Wave One (9 months)*
- *Report on Pre-piloting, Piloting and Dress Rehearsal Phases of the Infant Cohort at Wave One (9 months)*

#### 4.5.1.2 Available GUI data

Researchers can request to access multi-wave survey data from the GUI Cohort '08 at ages 9 months, 3 years, 5 years, 7/8 years, and 9 years, and for the GUI Cohort '98 at ages 9, 13, 17/18, and 20 years. Data from the COVID-19 Web Survey were released in autumn 2021. The survey covers 12/13-year-olds and primary caregivers from GUI Cohort '08 and 22-year-olds from GUI Cohort '98. The COVID-19 survey was designed to allow for data to be matched via ID code to previous cohorts in order to facilitate research on the effects of the pandemic on the study participants (Kelly *et al.*, 2021). Furthermore, anonymised data from 120 qualitative interviews collected during Wave 1 of both GUI Cohort '08 and GUI Cohort '98 are available for analysis from the Irish Qualitative Data Archive at Maynooth University.

#### 4.5.1.3 Accessing GUI data

There are two types of GUI microdata files available to access from each wave of the GUI data collection (Department of Children, Equality, Disability, Integration and Youth, 2020a). These are the Anonymised Microdata Files (AMFs) and Researcher

---

[30] GUI technical and data documentation can be found at https://www.growingup.ie/data-documentation/.

Microdata Files (RMFs).[31] AMFs include a basic set of variables approved by the CSO for distribution by the Irish Social Science Data Archive (ISSDA) and are available to all researchers once they meet the criteria for access. An access request form is available online via the ISSDA, and a signed copy must be submitted for approval. Researchers must provide a brief study outline, clarify the number of waves of data they require, and the required length of time for data use. Approval is granted within 7 working days of application (Department of Children, Equality, Disability, Integration and Youth, 2020b). Files are in SPSS, Stata or SAS format. From 2010 to 2019, the total number of successful AMF data access requests to ISSDA was 2,284 (Department of Children, Equality, Disability, Integration and Youth, 2020b).

RMFs are de-identified microdata files with a more detailed set of variables. These are available from the CSO to researchers at registered research organisations in cases where the AMF datasets do not meet the requirements of the research on their own. There is a stricter set of conditions to access the RMFs, and applications are made directly to the CSO using an RMF application form on the CSO website. A detailed proposal, including the study purpose, is necessary for application, and details of this can be found on the CSO's website.[32] Approval is granted within 6–8 weeks of application. To access the data, successful candidates must attend a data usage training course and sign a standard agreement for the duration of data access. If access is approved, researchers are then appointed as Officers of Statistics for a maximum of 1 year. Data analysis can only be conducted via a secure CSO Researcher Data Portal subject to two-factor authentication login procedures. Outputs will be released to researchers once they have been sanctioned as non-disclosive by the relevant CSO Statistician.[33] Ensuring that data are non-disclosive is considered very important. There are no fees associated with applying for or accessing any GUI data (see Appendix 4A for a comparison of selected birth cohort case studies by data accessibility and availability). Between 2010 and 2019, the total

---

[31] For further information about the GUI microdata files and how to access them, visit https://www.growingup.ie/information-for-researchers/.

[32] For more details on the RMF application procedure, see: https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/dataforresearchers/rmfapplicationprocedure/.

[33] For details of the GUI RMF data use instructions, see: https://www.cso.ie/en/media/csoie/aboutus-new/dataforresearchers/Instructions_on_the_use_of_GUI_data_for_Researchers.pdf.

number of successful RMF data access requests to ISSDA was 188 (Department of Children, Equality, Disability, Integration and Youth, 2020b).

GUI qualitative data can be accessed via the Irish Qualitative Data Archive at Maynooth University. Applications for access to these data can be made directly through the Irish Qualitative Data Archive website.

#### 4.5.1.4 GUI data access promotion

The GUI study engages with participants, data users and the public via its website and social media. The GUI website[34] has separate sections for study participants and researchers. The former focuses on informing participants about the study's progress, publications, and GUI contact details, while the latter publicises events, new data releases, data access and publications.

GUI has hosted a 1-day GUI research conference every year since 2009.[35] Researchers using GUI data are invited to submit abstracts and the event features approximately 25 presentations from early career researchers and more senior academics. To support the use of GUI data, the study team regularly hosts data workshops for interested researchers which provide an overview of the data alongside information on weighting, access and analysis. Workshop materials become available on the GUI website afterwards. Webinar recordings are also available on YouTube. GUI also offers infographic summaries for its reports which are available on the study's website.[36]

Established in 2013, the GUI Twitter account (@GrowingUpIre) communicates findings produced by the GUI study team and other researchers in addition to other GUI-related news (e.g. conferences, data workshops). As of February 2022, the account had more than 2,600 followers. This compares favourably with other "Growing Up" studies, such as Growing Up in Australia: The LSAC (@AIFS_LSAC with just over 400 followers) and Growing Up in New Zealand (@GrowingUpinNZ with more than 800 followers) but it has fewer followers than Growing Up in Scotland

---

[34] https://www.growingup.ie/

[35] See https://www.growingup.ie/information-for-researchers/gui-conferences/

[36] GUI infographic summaries can be found at https://www.growingup.ie/growing-up-in-ireland-publications/infographics/.

(@growingupinscot with more than 5,000 followers). Meanwhile, the MCS, DNBC and ELFE do not have Twitter accounts. Similarly, GUI is the only one of the named case studies in this report that has an Instagram account (with more than 1,000 followers).

## 4.5.2 MCS data access

### 4.5.2.1 MCS technical documents

There is a full and robust repository of technical documents available via the UK CLS's website. These include survey documentation such as technical reports, show cards, ethical review and consent guides, laboratory procedure guides, and interview instructions. In addition, there is a range of user guides for researchers working with the MCS data. These guides include data notes and reports, coding frames and derived variables guides. Full versions of the MCS questionnaires and various data collection tools are also available. For example, Table 4.13 lists technical documents that are available for MCS Sweep 1 (age 9 months).

**Table 4.13** Technical documents available for MCS Sweep 1 (age 9 months)

| **User guides:** |
|---|
| *User Guide for MCS Sweeps 1-5 of the Millennium Cohort Study* |
| *Longitudinal Family File Guide* |
| *MCS Data Handling Guide* |
| *MCS 9 Months-Age 11 Guide to the Datasets (Eighth Edition)* |
| *MCS 9 Months User Guide to Initial Findings* |
| *MCS 9 Months Guide to the SPSS Dataset* |
| **Questionnaires:** |
| *MCS 9 Months CAPI Questionnaire Documentation (2006)* |
| **Technical reports:** |
| *MCS 9 Months Technical Report on Sampling (4th Edition)* |
| *MCS First Survey: Technical Report on Instrument Development and Fieldwork* |
| **Data notes:** |
| *MCS Geographical Identifiers* |
| **Additional files:** |
| *MCS 9 Months ESRC End of Award Report* |
| *MCS 9 Months Codebook and Edit Instructions* |
| *MCS Ethical review and consent* |

#### 4.5.2.2 Available MCS data

Researchers can request access to currently available multi-wave MCS survey data from seven data sweeps: 2001 (9 months); 2004 (3 years); 2006 (5 years); 2008 (7 years); 2012 (11 years); 2015 (14 years); and 2018 (17 years). In addition, survey data are available from four MCS sub-studies; the COVID-19 survey (2021), Childcare in the Millennium Cohort Study (2004–2006), Survey of mothers who received assisted fertility treatment (2003), and Survey of health visitors (2002). There are separate survey datasets for each sweep and sub-study. Some data variables are comparable across datasets, while other variables are singularly associated with one data sweep. To access genetic and biomedical data related to the MCS, a written request must be made to the CLS Data Access Committee.[37] Data from oral fluid samples are available from the UK Data Service. Milk teeth are stored at the Institute for Child Health. DNA extracted from saliva biological samples are held at University of Bristol. To access genetic data associated with survey data, a specific data access agreement must be signed with the CLS Data Access Committee to create bespoke survey datasets (see Appendix 4A for a comparison of selected birth cohort case studies by available data).

#### 4.5.2.3 Accessing MCS data

Anonymised MCS data are freely accessible for researchers and can be found via the UK Data Archive which is run by the University of Essex.[38] MCS data fall under 'safeguarded' access: an approval process must be completed to access the data. This includes registering for a UK Data Service account, providing a brief outline of the research study, and accepting the terms and conditions of data use. Microdata can then be downloaded in multiple formats (i.e. SPSS or Stata files). Detailed information about MCS data access is available on the study website, including various details about the datasets (e.g. variables and sample sizes, details of regional levels of datasets, coding schemes, naming conventions, and availability of future data sweeps). There are no costs associated with accessing MCS microdata for research purposes.

---

[37] To access genetic and biomedical data related to the MCS, a written request must be made to the CLS Data Access Committee; see https://cls.ucl.ac.uk/data-access-training.

[38] UK Data Archive website: https://www.data-archive.ac.uk/.

#### 4.5.2.4 MCS data access promotion

The MCS does not have a stand-alone website, but it is featured on the CLS website, alongside three other British birth cohort studies (the 1958 National Child Development Study, the 1970 British Cohort Study, and Next Steps – the Longitudinal Study of Young People in England). The MCS page includes links to detailed study features, publications, data access and other content. The MCS does not have its own Twitter account, but the CLS account (@CLScohorts, with 6,000 followers, established in 2010) communicates MCS news and research findings. The CLS supports new and existing users of MCS data via training events and webinars publicised on its website. Similar to GUI, webinar recordings are available on YouTube (see Section 4.5.1.4). The CLS also runs a seminar series to communicate research findings and publicise its cohort studies, including the MCS.

### 4.5.3 DNBC data access

#### 4.5.3.1 DNBC technical documents

DNBC codebooks are available online for all data sweeps in Danish and English, with the exception of the Maternal Follow-up sweep. Interview guides are also available for each dataset. There are a handful of key papers composed by the study team which give context to the research. These are listed in Table 4.14.

**Table 4.14** Key DNBC study design papers

| **Background reports** |
|---|
| • Olsen, J. and Meder, K. (2014) 'Better health for mother and child – The Danish National Birth Cohort (DNBC), its structure, history and aims', *Norsk Epidemiologi*, *24*(1-2), 37-38. https://doi.org/10.5324/nje.v24i1-2.1756 |
| • Olsen, J. (2012) 'Nine months that last a lifetime. Experience from the Danish National Birth Cohort and lessons learned', *International Journal of Hygiene and Environmental Health*, 215(2), pp. 142–144. https://doi.org/10.1016/j.ijheh.2011.10.015 |
| **Methodological design reports** |
| • Greene, N., Greenland, S., Olsen, J. and Nohr, E. A. (2011) 'Estimating bias from loss to follow-up in the Danish National Birth Cohort', *Epidemiology*, 22(6), pp. 815–822. https://doi.org/10.1097/EDE.0b013e31822939fd |

- Bliddal, M., Liew, Z., Pottegård, A., Kirkegaard, H., Olsen, J. and Nohr, E. A. (2018) 'Examining Nonparticipation in the maternal follow-up within the Danish national birth cohort', *American Journal of Epidemiology*, 187(7), pp. 1511–1519. https://doi.org/10.1093/aje/kwy002
- Jacobsen, T. N., Nohr, E. A. and Frydenberg, M. (2010) 'Selection by socioeconomic factors into the Danish National Birth Cohort', *European Journal of Epidemiology*, 25(5), pp. 349–355. https://doi.org/10.1007/s10654-010-9448-2
- Nohr, E. A., Frydenberg, M., Henriksen, T. B. and Olsen, J. (2006) 'Does low participation in cohort studies induce bias?', *Epidemiology*, 17(4), pp. 413–418. https://doi.org/10.1097/01.ede.0000220549.14177.60

### 4.5.3.2 Available DNBC data

Researchers can request access to survey data from the following data sweeps: History of Births; Miscarriage Project; DNBC Interviews 1–4; Food Frequency Questionnaire (FFQ); Lifestyle During Pregnancy; 7-year Follow-up; 11-year Follow-up; Follow-up among mothers; Dietary Habits of 14-year-olds; Puberty Follow-up; and COVID-19. Survey data for each sweep are available in separate data files. Dataset requirements must be confirmed with the DNBC data managers who then construct bespoke datasets as per individual research study requirements. This process means that DNBC datasets are not accessible in full; instead, selected variables are used as per the bespoke requirements of each secondary research study. For example, regarding DNBC interviews 1–4, each interview sweep (1–4) is contained in a separate dataset. Biological samples, which are stored in the DNBC biobank, are also available for research purposes. However, data access to biological samples must meet ethical and legal compliance with Danish law. Overall, data access is limited, with a range of regulations in place to access the data. Access to utilise these data must be sought from three groups: the Scientific Ethical Committee, the Danish Data Protection Agency, and the DNBC management team.[39] Samples were collected from 100,000 pregnant women at the start of or at later stages of pregnancy, and approximately 60,000 umbilical cord samples were collected as well.[40] The total number of biobank samples is 625,614. These include:

---

[39] Further information on how to apply to utilise biological samples data can be found at https://www.dnbc.dk/access-to-dnbc-data.

[40] For sample collection information regarding the DNBC, visit www.eithealth-scandinavia.eu.

- 192,834 buffy coat samples (a concentration of the white blood cells and platelets in a blood sample)
- 3,593 DNA samples
- 205,931 filter paper samples (small volumes of blood collected on filter paper)
- 223,256 plasma samples

### 4.5.3.3 Accessing DNBC data

The DNBC data are open to researchers and research studies who will undertake work that adheres to the policy and overall aim of the DNBC. The data are stored on a server at Aarhus University and are only accessible via a virtual desktop infrastructure (VDI) secure connection to the server.[41] To access data, applicants must complete an application form and write a short protocol of the project using a set template. The application pack must be submitted to the DNBC Steering Committee via direct email. It takes 6–8 weeks to process the application. In addition, applicants must receive permission to use the data from the Danish Data Protection Agency. Regarding international research studies, as the DNBC is funded by Danish funders, the DNBC team prefers that data analysis work is conducted in Denmark in collaboration with Danish researchers.[42] The CSO in Ireland has a similar but more defined policy regarding GUI RMF data; this policy states that requests to access RMFs will only be granted to researchers physically located to conduct analysis in the Republic of Ireland.[43] See Appendix 4A for a comparison of selected birth cohort case studies' approval waiting times.

Since 1 October 2018, administrative fees apply to process data access requests, and for the construction of the bespoke datasets by DNBC data managers (see Table 4.15). Additional annual fees are also charged for connection to the DNBC server and storage space. Data are then accessed via secure connection to the Aarhus University server (see data access information above). See Appendix 4A for a comparison of selected birth cohort case studies' data access fees.

---

[41] For further information about accessing DNBC data, visit https://www.dnbc.dk/access-to-dnbc-data.

[42] For further information regarding DNBC open access policy and data use, visit https://www.dnbc.dk.

[43] For more detail about the CSO Ireland policy on access to RMFs, visit https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/dataforresearchers/policies/#:~:text=Access%20to%20RMFs%20is%20a,such%20access%20may%20be%20granted.&text=i)%20the%20intended%20results%20of,be%20published%20or%20otherwise%20disseminated.

**Table 4.15** Breakdown of the DNBC survey data access fees[44]

| |
|---|
| • Application processing fee: DKK 4,000 (Danish krone)/EUR 537.90 (euro) |
| • Data access for projects with one intended scientific publication: DKK 21,000/EUR 2,823.99 |
| • Data access for projects with up to four intended scientific publications: DKK 41,000/EUR 5,513.51 |
| • VDI server fees (annual fee): DKK 5,000/EUR 672.35 |
| • Construction of dataset (approximate): DKK 10,000–15,000/EUR 1,344.79–2,017.19 |

Several costs are associated with biological samples. These are divided into categories based on the number of samples (i.e. less or more than 500 samples; samples retrieved from manual or automated storage; sample aliquoting; DNA extraction; laboratory services; and IT services).

#### 4.5.3.4 DNBC data access promotion

To the authors' knowledge, there is little publicly available information regarding the types of training, workshops or conferences associated with the DNBC data. In May 2019, a scientific symposium was held in Denmark to mark 20 years of the DNBC study. At the event, presentations related to DNBC data were given.

### 4.5.4 ELFE data access

#### 4.5.4.1 ELFE technical documents

The ELFE study technical documentation is available on the study website. Questionnaires, protocols, interview guides, self-administered questionnaire templates, and weighting guides are available under the 'research' tab of the study website. All documents are available in English. A full list of available technical documents can be found in Table 4.16.

---

[44] A breakdown of the DNBC survey data access fees can be found at https://www.dnbc.dk/access-to-dnbc-data.

**Table 4.16** Questionnaires and technical documents as recorded on the ELFE website

| | |
|---|---|
| Maternity unit survey | Maternity survey protocol<br>Maternity mother interview<br>Mother's self-administered questionnaire<br>Maternity medical record |
| Survey at 2 months | 2 months survey protocol<br>Mother 2 months interview<br>Father 2 months interview |
| Diet between 2 and 10 months | Infant's diet 2-6 months questionnaire<br>Infant's diet 6-10 months questionnaire |
| Survey at 1 year | 1 year survey protocol<br>Mother 1 year interview<br>Father 1 year interview |
| Survey at 2 years | 2 years survey protocol<br>Mother 2-year interview<br>Father 2-year interview |
| Survey at 3 years | 3 years survey protocol<br>3 years survey interview |
| Weighting | Weighting inclusion<br>Weighting 2 months<br>Weighting ELFE surveys general document |

Source: https://www.elfe-france.fr/en/the-research/access-to-data-and-questionnaires/

### 4.5.4.2 Available ELFE data

Datasets are currently available for the following surveys: Maternity unit survey; Survey at age 2 months; Diet between ages 2 and 10 months; Survey at age 1 year; Survey at age 2 years; Survey at age 3.5 years; Preschool survey; and ELFE et services de Protection Maternelle et Infantile (ELFE-PMI) survey. Each sweep is a separate data file. Like the DNBC study, it appears that datasets are not assessable in full; instead, selected variables are used as per the bespoke requirements of each secondary research study.

### 4.5.4.3 Accessing ELFE data

The ELFE study utilises an open access data policy for researchers under set privacy and security conditions as listed in the data access charter document published on its website.[45]This document is accessible only in PDF format, and only in French. Access

---

[45] www.Elfe-france.fr

requests must be made via a secure platform that details the data accessible. To process a data request, research teams must define their project aims, provide a description of the research team and collaborators, and select the required ELFE data variables for their study. Priority is given to ELFE study research teams for the initial 18 months after data are made available. Following this, data access requests are opened to other researchers. Data access requests are reviewed by the ELFE data access committee (Comité d'Accès aux Données Elfe; CADE). Successful selection requirements include whether the proposals are eligible for use, the scientific relevance of proposed studies to the ELFE research objectives, well-considered participant privacy protocols, and whether the secondary research will be in the public interest. Researchers who are not affiliated with the primary ELFE study also need to gain study approval from the French National Data Protection Authority (Commission Nationale de l'Informatique et des Libertés; CNIL) and include a statement of intent to commit to the ELFE study methodological guidelines.

#### 4.5.4.4 ELFE data access promotion

There is minimal information available regarding the types of training, workshops, or conferences for researchers to maximise their use of ELFE birth cohort data. However, there are plenty of resources that are used by the ELFE team to engage with cohort participants, their families, and the public. Media releases were used between 2011 and 2018 to inform the public and media of key milestones in the study. The study also has a newsletter subscription for news and updates related to the study. For the cohort families, an annual video conference is hosted by the ELFE team to engage parents in the study, inform them about future surveys, and present the latest research findings. In addition, there is a section of the ELFE website dedicated to video updates by the ELFE research team and collaborators who use video stories to describe their work and talk about their cutting-edge findings. Video topics include an introduction to the ELFE study, what happens at a home visit when the child participant is aged 3.5 years, a 'vox pop' (i.e. popular opinion insights via comments from members of the public) about the questions families might have, why child cohorts are essential research data collection designs, what the ELFE study contributes to the discipline of social science, and examples of ELFE research in the disciplines of health and the environment[46] (see Appendix 4A for a comparison of selected birth cohort case studies' mechanisms for maximising data).

### 4.5.5 GenV data access

As the GenV study has just started data collection, there is limited information available regarding data access processes, available data, or mechanisms for maximising the data for the study. In addition, there is not yet a body of published research to identify. However, during the period 2022–2023, GenV is planning its data access and biosample access procedures. As stated on the study website, data access for the study will work under the principles of equal access underpinned by the findable, accessible, interoperable, and reusable (FAIR) principles and the Five Safes framework. The Five Safes framework offers an outline for evaluating and managing disclosure risk related to data use and data release (Ritchie, 2017). The five elements of the Five Safes framework are: safe people, safe projects, safe settings, safe data, and safe outputs. Often used by statistical data agencies, including the CSO (Ireland) and the Office of National Statistics (UK), and across the social sciences, the framework is now also utilised by government bodies, the health sector, and private institutions. The FAIR principles are a guidance framework for ethically sharing data in such a way as to maximise its use (Wilkinson *et al.*, 2016).[47] The GenV research team has also clarified via the study website that there will be phases of "protected data access" after which data will be made available in full, but no details on the length of this period are available on the website.[48] Currently, there are a handful of published papers which give context to the research by the study team members. These include a summary of methods for integrating registry data into GenV, by Sung *et al.* (2021); a discussion of broad principles for incorporating randomised controlled trials into GenV by Wake *et al.* (2020); and a review of conceptual frameworks for inclusion in GenV by Wang *et al.* (2021).

---

[46] ELFE study video updates can be found on the study website at https://elfe-france.fr/en/document-library/videos/institutional-films/.

[47] The FAIR principles can be found at https://www.go-fair.org/fair-principles/.

[48] https://www.genv.org.au/for-researchers/the-benefits-of-the-genv-data-repository/

## 4.6 Policy use of birth cohort study analyses

Birth cohort studies have significant potential to influence public policy because they follow large groups of children and families over time (Power, Kuh & Morton, 2013). The examples below offer insight into the wide range of policy areas that birth cohort data can inform. For conciseness, these are limited to GUI, MCS, ELFE and DNBC.

According to an analysis of applications for data access during the period 2010–2019, GUI study data have been used by 24 Government Departments and public bodies in Ireland to develop knowledge and influence policy changes (Department of Children, Equality, Disability, Integration and Youth, 2020b). For example, the Department of Health has used GUI data in the Healthy Ireland policy framework, and the Department of Education has used GUI data in its review of career guidance in an Irish context (Department of Children, Equality, Disability, Integration and Youth, 2020b). GUI data have frequently been used by DCEDIY to influence policy development and service provision. Examples include the use of data in two national strategies: *Better Outcomes, Brighter Futures: The national policy framework for children & young people 2014 – 2020* and *First 5: A Whole-of-Government Strategy for Babies, Young Children and their Families 2019-2028* (Department of Children, Equality, Disability, Integration and Youth, 2020b).Given that GUI contains data on sexual orientation and gender identity, the study is also relevant to the *LGBTI+ National Youth Strategy: 2018-2020 established* by DCEDIY (Department of Children, Equality, Disability, Integration and Youth, 2020b). In addition, GUI data were used to support the work undertaken by the Centre for Effective Services, which conducted an evaluation of the Area-Based Childhood (ABC) Programme, an early intervention and prevention initiative funded by DCEDIY (Department of Children, Equality, Disability, Integration and Youth, 2020b). Meanwhile, analyses of longitudinal GUI data by the GUI study team and researchers at the Economic and Social Research Institute (ESRI) have informed policy-makers' understanding of health and well-being in childhood and adolescence (Nolan & Smyth, 2021), children's transition to primary school (Smyth, 2018), the home learning environment, parental and non-parental childcare, and children's cognitive outcomes at ages 3 and 5 years (McGinnity *et al.*, 2017). More recently, data from the special COVID-19 web survey provided unique insights into schoolchildren's experiences during the first year of the pandemic (Murray *et al.*, 2021).

Similarly, the MCS has informed policy, practice and public debate in a range of areas, including parenting, poverty, child development, and health (Johnson & Antill, 2011). For example, MCS data identified that 12% of cohort children were unimmunised or did not receive full immunisation, which was referenced by National Institute for Health and Care Excellence 2009 guidelines (UCL, 2014). MCS data were also used in the evaluation of the UK Children's Fund and Sure Start national programmes (IoE, 2010). A recent strategic review of UK longitudinal studies by Davis-Kean *et al.* (2018) included several policy impact case studies where study findings informed policy, including MCS research on the benefits of breastfeeding, the consequences of household poverty, and the prevalence of overweight and obesity. Although it can be difficult to attribute specific policy changes to MCS findings, the review found that UK government analysts tended to be "overwhelmingly positive about the value of longitudinal data in informing policy" (Davis-Kean *et al.*, 2018, p. 3). As the MCS cohort are now in the third decade of their lives, the study holds vast potential to understand the importance of earlier experiences in transitions to adulthood (Joshi & Fitzsimons, 2016).

The ELFE study data have also been used by researchers to inform policy and practice. On the 10th anniversary of the study in 2021, the ELFE study team documented 10 case studies of novel research findings of relevance to public debates and policy.[49] For example, the French government extended the national paternity leave in 2021 from 14 to 28 days after ELFE findings showed that paternity leave was key to father–child bonding and more equitable gender division of labour in the home (Pailhé, Solaz & Tô, 2018). Other influential findings summarised on the ELFE website include: the harmful effects of pesticides in pregnancy; the link between prolonged screen time use at ages 2 and 3 years and higher risks of behavioural problems; the lack of evidence that hypoallergenic infant formula protected infants from developing allergies; the importance of health screening in preschool; the health benefits of breastfeeding for infants; the positive effects of centre-based care

---

[49] https://www.elfe-france.fr/en/results-so-far/10-years/

for early language development; and the experiences of children and families during the first wave of the COVID-19 pandemic. Meanwhile, DNBC has produced valuable epidemiological evidence about different pregnancy risks and adverse pregnancy outcomes. For example, DNBC findings informed the Danish Institute of Medicine's new recommendations for weight gain in pregnancy and other antenatal care policies (Nybo Andersen & Olsen, 2011).

## 4.7 Conclusion

This chapter discussed the main types of primary and secondary analyses produced by the GUI study and the named cohort studies (MCS, DNBC, ELFE, and GenV), and outlined the data access systems and promotion activities of each of those studies (where information was available). In the conclusion, we briefly discuss best practices for facilitating data access and maximising data use for research.

### 4.7.1 Facilitating data access

Birth cohort studies tend to share their anonymised data with researchers, but modalities for data access and conditions of data use vary substantially. Some studies require a detailed research proposal and take several weeks to review. For example, an application for DNBC data can take 6–8 weeks to review.[50] Others ask for only a brief description of the proposed analysis and notify applicants promptly. For example, applications for GUI AMFs are processed within 1 week from receipt of application (Department of Children, Equality, Disability, Integration and Youth, 2020b). Of the four named cohort studies, DNBC is the only one that charges researchers fees for accessing the data. Comprehensive detail is available online for the named case studies, with guidance on how to apply for data. GUI's data access procedure is two-tiered; while a large subset of the survey questions are available as AMFs to researchers via a quick approval process (within 1 week), applications for a more detailed set of variables in the RMFs (for researchers who have exhausted the use of AMF datasets) are subject to stringent conditions of use and take 4–8 weeks to process (Department of Children, Equality, Disability, Integration and Youth,

---

[50] https://www.dnbc.dk/access-to-dnbc-data

2020b). Once access is approved, RMF data can only be analysed on a secure remote platform.

### 4.7.2 Mechanisms for maximising data use

There are different ways to maximise data use for research, including, for example, training events, seminars/conferences, and online visibility. Training events can address different needs by supporting data access (i.e. how to obtain microdata files for research) and data use (e.g. survey design, weighting, data linkages, and data analysis). Research seminars and conferences promote the survey by disseminating and showcasing its findings. Online visibility, including on social media, is key to raising awareness of the study and promoting it to diverse audiences (e.g. early career researchers).

The four birth cohort case studies discussed in this chapter vary in the extent to which they leverage these resources. GUI has a strong social media presence, an informative website, an annual research conference, and data user workshops. The MCS benefits from the resources and online presence of the CLS, but it might benefit from cultivating a separate modern identity from the older UK birth cohorts. In contrast, the DNBC and ELFE have websites, but no apparent social media presence. There is no information on those websites about data user workshops or research seminars.

Studies need to develop evidence-based strategies for maximising the use of their data and periodically evaluating their success. In addition to the mechanisms listed above, studies can take part in international research collaborations and facilitate transnational access visits to research institutions (see, for example, the COhort cOmmunity Research and Development Infrastructure Network for Access Throughout Europe (COORDINATE) Network website).[51] Given the complex nature of longitudinal data, training in advanced data analysis methods needs to be enabled (Johnson & Antill, 2011). For example, studies could provide information about relevant statistical training courses on their websites. This should include not only advanced statistical techniques for panel data analysis, but also training on data

---

[51] Information on international research collaborations and transnational access visits can be found at www.coordinate-network.eu/.

linkage (e.g. how to deal with linkage error) and biomedical data analysis, where applicable. This could be part of a broader national strategy of researcher training. A strategic review of UK longitudinal studies by Davis-Kean *et al.* (2018) recommended that the ESRC commission a review of the provision of longitudinal data training and ensure sufficient academic training, especially at graduate and postgraduate levels.

# Appendices

## Appendix 4A

Comparison of selected birth cohort case studies by data accessibility, availability, documentation, and data maximisation.

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) |
|---|---|---|---|---|
| **Data access formats** | Anonymised Microdata Files (AMFs)<br>Researcher Microdata Files (RMFs) | Anonymised SPSS or Stata data files<br>Genetic/biomedical data | Bespoke microdata sets<br>Biobank samples:<br>• 192,834 buffy coat samples<br>• 3,593 DNA samples<br>• 205,931 filter paper samples<br>• 223,256 plasma samples | The ELFE study utilises an open access data policy for researchers under set privacy and security conditions as listed in the data access charter document published on its website (www.Elfe-france.fr). This document is accessible only in PDF format, and only in French. |
| **Data access process** | **AMF:** access request form; research proposal<br>**RMF:** attendance to terms of data usage training; sign an agreement contract for the duration of data access | Brief outline of the research study; acceptance of terms and conditions of data use.<br>Download microdata via a researcher's UK Data Service account.<br>Genetic and biomedical data: via written request to CLS Data Access Committee.<br>To access genetic data associated with survey data, a specific data access agreement must be signed to create individual survey datasets. | Application form: short protocol of the project using a set template submitted to the DNBC Steering Committee via direct email.<br>In addition, permission from the Danish Data Protection Agency.<br>Priority in accessing data is given to researchers working on already established projects or to researchers tied into pre-existing data agreements. | Access requests via a secure platform; include project aims, a description of the research team and collaborators, and the required ELFE data variables for study.<br>Priority is given to ELFE study research teams for the initial 18 months after data are made available. After 18 months, data are opened to other researchers.<br>Researchers who are not affiliated with the primary ELFE study also need to gain study approval from the CNIL and include a statement of intent to commit to the ELFE study methodological guidelines. |

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) |
|---|---|---|---|---|
| **Data access fees** | ✕ | ✕ | **Administrative fees**<br>Application processing fee: DKK 4,000/EUR 537.90<br>Data access for projects with one intended scientific publication: DKK 21,000/EUR 2,823.99<br>Data access for projects with up to four intended scientific publications: DKK 41,000/EUR 5,513.51<br>VDI server fees (annual fee): DKK 5,000/EUR 672.35<br>Construction of dataset (approximate): DKK 10,000–15,000/EUR 1,344.79–2,017.19<br>**Costs for biological samples**<br>Based on the number of samples (i.e. less or more than 500 samples), samples retrieved from manual or automated storage, sample aliquoting, DNA extraction, laboratory services, and IT services. | ✕ |
| **Approval time** | **AMF:** within 7 working days of application<br>**RMF:** within 4–8 weeks of application | Immediate approval upon authenticated registration and accepting the terms and conditions of use. | Unclear | Unclear |

| | GUI Cohort ’08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) |
|---|---|---|---|---|
| **Available data** | **Survey data from GUI Cohort ’08**<br>Ages 9 months; 3 years; 5 years; 7/8 years; 9 years<br>**Survey data from GUI Cohort ’98**<br>Ages 9 years; 13 years; 17/18 years<br>**COVID-19 survey data**<br>• For 12/13-year-olds and primary caregivers from Cohort ’08<br>• For 22-year-olds from Cohort ’98<br>**Qualitative data**<br>120 interviews collected during Wave 1 of both **GUI Cohort ’08** and **GUI Cohort ’98.** | **Seven data sweeps**<br>2001 (9 months); 2004 (3 years); 2006 (5 years); 2008 (7 years); 2012 (11 years); 2015 (14 years); and 2018 (17 years)<br>**Four MCS sub-studies**<br>COVID-19 survey (2021); Childcare in the Millennium Cohort Study (2004–2006); Survey of mothers who received assisted fertility treatment (2003); Survey of health visitors (2002) | **DNBC data sweeps**<br>History of Births; Miscarriage Project; DNBC Interviews 1–4; FFQ; Lifestyle during pregnancy; 7-year Follow-up; 11-year Follow-up; Follow-up among mothers; Dietary Habits of 14-year-olds; Puberty Follow-up; COVID-19.<br>DNBC datasets are not accessible in full. Instead, selected variables are used as per the bespoke requirements of each secondary research study. | **ELFE surveys**<br>Maternity unit survey; Survey at 2 months; Diet between 2 and 10 months; Survey at 1 year; Survey at 2 years; Survey at 3.5 years; Preschool survey; ELFE-PMI survey.<br>Datasets are not assessable in full. Instead, selected variables are used as per the bespoke requirements of each secondary research study. |
| **Technical reports** | • *Cohort ’08 at 9 months Codebook for Wave 1 of the Infant Cohort*<br>• *Data Dictionary for Wave 1 of the Infant Cohort (at 9 months)*<br>• *Questionnaires for Wave 1 of the Infant Cohort (at 9 months)*<br>• *Sample design and response in Wave 1 of the Infant Cohort (at 9 months)*<br>• *Summary Data Dictionary for Wave 1 of the Infant Cohort (at 9 months)*<br>• *Summary Guide to Wave 1 of the Infant Cohort (at 9 months)*<br>• *Variable Naming Conventions and Longitudinal Data Dictionary for Wave 1 and Wave 2 of Cohort ’08* | **User guides:**<br>• *User Guide for MCS Sweeps 1-5 of the Millennium Cohort Study*<br>• *Longitudinal Family File Guide*<br>• *MCS Data Handling Guide*<br>• *MCS 9 Months-Age 11 Guide to the Datasets (Eighth Edition)*<br>• *MCS 9 Months User Guide to Initial Findings*<br>• *MCS 9 Months Guide to the SPSS Dataset*<br>• *MCS 9 Months-Age 11 Guide to psychological and developmental inventories* | Limited access to technical documents and key finding reports and methodological design reports.<br>Codebooks are available online for all data sweeps in Danish and English except for the Maternal Follow-up sweep.<br>Interview guides are available for each dataset. | • Questionnaires<br>• Protocols<br>• Interview guides<br>• Self-administered questionnaire template<br>• Weighting guides<br>All documents are available in English. |

| | GUI Cohort '08 (Ireland) | MCS (UK) | DNBC (Denmark) | ELFE (France) |
|---|---|---|---|---|
| | | **Questionnaires:**<br>*MCS 9 Months CAPI Questionnaire Documentation (2006)*<br>**Technical reports:**<br>• *MCS 9 Months Technical Report on Sampling (4th Edition)*<br>• *MCS First Survey: Technical Report on Instrument Development and Fieldwork*<br>**Data notes:**<br>*MCS Geographical Identifiers*<br>**Additional files:**<br>• *MCS 9 Months ESRC End of Award Report*<br>• *MCS 9 Months Codebook and Edit Instructions*<br>• *MCS Ethical review and consent* | | |
| **Mechanisms for maximising the data** | • Study website<br>• Annual 1-day conference (free)<br>• Data workshops<br>• Twitter | The MCS does not have a stand-alone website, but it is featured on the CLS website (https://cls.ucl.ac.uk/).<br>MCS information shared via CLS Twitter account (@CLScohorts)<br>CLS supports users of MCS data via training events and webinars publicised on its website.<br>Webinar recordings on YouTube.<br>CLS runs a seminar series to communicate research findings and publicise its cohort studies, including the MCS. | Study website<br>Minimal information available regarding the types of training, workshops or conferences available | Study website.<br>Video updates by ELFE research team and collaborators.<br>Minimal information available regarding the types of training, workshops, or conferences for researchers to maximise use of cohort data.<br>Numerous resources used by the ELFE team to engage with cohort participants, their families, and the public, including media releases, newsletter subscription, and annual video conference for cohort families. |

# Appendix 4B

## Web of Science search

| Case study name | Search return number |
|---|---|
| GUI | 167 |
| MCS | 358 |
| DNBC | 257 |
| ELFE | 192 |

# Chapter 5: Stakeholder consultation

## 5.1 Introduction

In this chapter, we report the results of a stakeholder consultation that was carried out to identify the research needs for a potential new GUI birth cohort study. This consultation focused on the following: (1) key research and policy needs for each stakeholder; (2) stakeholder advice on relevant participant groups and survey content for a new GUI birth cohort; and (3) potential enablers for, and barriers to, accessing cohort study data in their organisations.

## 5.2 Methods

The consultation was approved by the UCD Human Research Ethics Committee through a low-risk application (HS-E-21_126-Symonds). Stakeholders gave their informed and voluntary written consent to participate in the consultation interviews.

Potential stakeholders were identified using various strategies, including Internet searches, government and university publications, and GUI data users listed in the ISSDA.[52] Stakeholders prioritised for recruitment were senior professionals representing a mixture of genders and types of organisations responsible for research, policy, and practice regarding child welfare and well-being.

Stakeholders within each organisational category were contacted one at a time until the target recruitment of 18 stakeholders was reached. The pattern of invitation acceptance resulted in a gender imbalance, with 14 females and 4 males agreeing to be interviewed. However, the spread across organisational sectors was relatively balanced, with four of the stakeholders working in Government Departments, four stakeholders working in State agencies, five working in non-governmental organisations (NGOs), and five working in higher education institutions (HEIs). No research institutes (e.g. ESRI) participated in this research.

Given the COVID-19 restrictions during the consultation period (September–December 2021), interviews with stakeholders took place using Zoom and were recorded via the application's built-in function and with a secondary audio recording

---

[52] https://www.ucd.ie/issda/data/growingupinirelandgui/guiregisterofuse/

device. Consent was mandatory for participation and all interviewees had to return the digitally signed forms before taking part in the interview. The Zoom interviews were video recorded using the programme's built-in function, which also automatically created audio files. The secondary audio recording device was used to create backup audio recordings of the Zoom interviews. Each interview lasted 50 minutes on average. Interviews were transcribed by the research team using Microsoft Word. All data were protected using secure file storage and transfer. Data were processed in NVivo Version 12 software, a specialist computer programme for analysing qualitative data.

To analyse the stakeholder consultation data, we first organised the data into major categories which aligned with the interview questions. Next, within each major category, we further grouped the data into minor categories representing what the stakeholders expressed as important to them. Finally, we organised the major categories into four broader themes. The remainder of this report presents the thematically organised findings from the interviews. The following summary of results is organised into four main parts: (1) stakeholders in context; (2) stakeholder needs; (3) stakeholder views on a new GUI birth cohort study; and (4) GUI data access. 5.

For an overview of the stakeholders' response patterns across key themes, see Table 5.1.

**Table 5.1** Density of coded text by participant group

| Major and minor themes | Percentage of segments coded (%) | | | | |
|---|---|---|---|---|---|
| | Government Departments | State agencies | NGO | HEI | Total |
| **Stakeholders in context** | | | | | |
| Stakeholders' prior engagement with GUI | 6.90 | 17.24 | 51.72 | 24.14 | 100.00 |
| Organisational activities | 16.42 | 25.37 | 46.27 | 11.94 | 100.00 |
| Changing landscapes and remits | 11.54 | 44.23 | 32.69 | 11.54 | 100.00 |
| **Stakeholder needs** | | | | | |
| Policy-related needs | 34.78 | 13.04 | 52.17 | 0.00 | 100.00 |
| Evidence needs | 12.90 | 22.58 | 37.10 | 27.42 | 100.00 |

| **Major and minor themes** | **Percentage of segments coded (%)** | | | | |
|---|---|---|---|---|---|
| | **Government Departments** | **State agencies** | **NGO** | **HEI** | **Total** |
| **Stakeholders' views on a new GUI birth cohort study** | | | | | |
| GUI participant groups | 40.58 | 4.35 | 44.93 | 10.14 | 100.00 |
| GUI content areas | 18.12 | 25.36 | 34.78 | 21.74 | 100.00 |
| GUI methodology | 12.87 | 36.63 | 24.75 | 25.74 | 100.00 |
| **GUI data access** | | | | | |
| Education and training in secondary data analysis | 10.00 | 40.00 | 50.00 | 0.00 | 100.00 |
| Different modes of accessing GUI data | 20.00 | 22.50 | 30.00 | 27.50 | 100.00 |
| Lack of organisational resources for data analysis | 11.11 | 40.74 | 37.04 | 11.11 | 100.00 |

## 5.3 Stakeholders in context

### 5.3.1 Stakeholders' prior engagement with GUI

*Summary:* Stakeholders had a range of levels of engagement with GUI within their organisations and professional backgrounds.

Generally, participants who had less experience of engaging with GUI mentioned having read the GUI reports, but had limited further engagement with the study. Other stakeholders had more experience of engaging with GUI, including involvement in the original study design and having commissioned research using GUI data. When describing their encounters with GUI, many stakeholders spoke of the utility of the study in relation to their organisational needs, and societal needs for evidence-informed practices for improving children's lives.

> *"The government can't answer its own question about better outcomes for children without the data and we can't advocate for children without that kind of data."* Stakeholder 5 (NGO)

### 5.3.2 Organisational activities

*Summary:* Stakeholders' organisations engaged in diverse activities that supported child welfare and well-being.

Stakeholders from Government Departments described their organisation as being diverse in scope and remit, depending on the Department's current statutory responsibilities. This group's commonly cited activities were policy-making and evaluation, policy monitoring, and service and programme creation and evaluation. All Government Department and State agency stakeholders indicated an understanding of the importance of working in an interconnected manner with other Government Departments on issues regarding child well-being.

> *"Our department is very broad-based. So, there's a huge focus on early years education and support…areas where there's child deprivation or abuse…adoption policy, youth services, family integration, refugees, and equality policies. So, the department has grown quite a lot."* Stakeholder 2 (Government Department)

Stakeholders from NGOs also engaged in diverse activities for supporting child and youth well-being, including providing legal services, training, policy evaluation, and political advocacy.

The final group were stakeholders from the higher education sector, all under contractual employment within an Irish HEI. For this group, common professional activities included creating new knowledge to inform various forms of policy-making.

### 5.3.3 Changing landscape and remits

*Summary:* Stakeholders worked within a changing policy and organisational landscape for supporting child welfare and well-being. This strengthened the argument for a new birth cohort study.

All participants reported that their organisation operated using strategies and policies that were regularly re-evaluated and updated (typically every 3–5 years). This was reflective both of organisational development and of the changing landscape in which organisations functioned. This situation helped clarify, for some stakeholders, the relevance of the GUI study.

> *"We absolutely need it [GUI]! Children's lives are very different from what they were 10 never mind 20 odd years ago. The country is fast changing so it is important to understand childhood as well as the changes in how children actually live and the kind of lives that they live."* Stakeholder 5 (NGO)

Another commonly discussed issue was the changing policy landscape in Ireland over the past decade. The policies discussed included national strategies such as those outlined in the First 5 whole-of-government strategy to improve the lives of babies, young children, and their families (Department of Children and Youth Affairs, 2018).

Stakeholders also discussed how GUI data could help track complex, longitudinal changes in children and youth's quality of life in relation to new policies and guidelines related to their specific remits.

> *"We talk of it in terms of determinants, the popular understanding of determinants is like one thing equals another, but one thing influences another or many things in our case influence another, it's complicated, but GUI helps us unpack that."* Stakeholder 8 (HEI)

## 5.4 Stakeholder needs

### 5.4.1 Policy-related needs

*Summary:* Stakeholders discussed needing better methods of evaluating policy, more integrated actions between organisations to affect policy, and a strong government commitment to evidence-informed policy on child welfare and well-being.

Stakeholders from Government Departments described the challenge of evaluating the impact and efficiency of complex and interacting policies. Here, stakeholders mentioned the importance of a potential new birth cohort study for better capturing policy-related changes.

> *"If it has a [negative] impact on something else it's important to know that too. Because that can often happen because we have so many schemes and the*

> *schemes are complicated, and they interact with each other."* Stakeholder 9 (Government Department)

Stakeholders from NGOs indicated challenges arising from a lack of clarity around remits and statutory responsibilities. For one stakeholder, this was experienced in the form of a lack of policy on specific groups of migrant children. The need expressed here was for more integrated actions between organisations to target specific challenges to child welfare.

> *"[We need] a childproofing of our immigration system…getting kids as quickly as possible to secure immigration status that they hold themselves – they're not just taken as being part of a family unit, and that comes into question when the family unit breaks down – to give them security and long-term immigration status [and] access to citizenship as quickly as possible."* Stakeholder 7 (NGO)

Finally, stakeholders from Government Departments, NGOs, and HEIs indicated a need for strong government policy commitments to support child welfare and well-being.

> *"We need policy that directs attention to children – so we need policies that promote more research, and that means more money, and more money needs to be spent proportionately on children's research. That [spending money on children's research] is at present the case, but we also need policy about how we use such information to implement strategies to benefit all children but also those children who are in most need of extra help."* Stakeholder 8 (HEI)

### 5.4.2 Evidence needs

*Summary:* Stakeholders discussed their needs for disaggregated socioeconomic and demographic data, national data linkage, and the public reuse of data on children held by private organisations and NGOs.

The most mentioned evidence needs among Government Departments, NGOs, and HEIs were around the need for disaggregated socioeconomic and demographic data for an accurate overview of Ireland's dynamic population trends.

> *"We need disaggregated data on the basis of ethnicity not just to inform policy and service provision, but also to identify gaps where they exist. Particularly in relation to minority ethnic groups to identify discrimination where it exists and identify next proactive steps to address that. We see data as fundamental to all*

> *of our work, but equally the state needs the data to inform policy and to report back to European Institutions and Human Rights monitoring body."* Stakeholder 7 (NGO)

Similar to points regarding data linkage in Chapter 2, stakeholders also discussed the need for such data linkage across organisations, Government Departments, and existing data management systems in organisations such as the HSE, Higher Education Authority, and the CSO to provide a more holistic overview of child welfare and well-being in Ireland.

> *"There is hardly any integration between different information management systems, so education, health, welfare data don't talk to each other…it is due to an underdeveloped tradition of using highly integrated data management systems."* Stakeholder 10 (HEI)

Stakeholders also discussed the use of data collected by private organisations and NGOs and indicated that these datasets lack visibility. Stakeholders signalled that these datasets could be pulled into policy directives and strategies like the *First 5* or the EU Roma Strategic Framework.

Stakeholders commented that, prior to the availability of GUI data in Ireland, they relied on international data sources and other available datasets to analyse national trends. Notwithstanding the subsequent establishment of the GUI study, stakeholders from Government Departments stressed the need for even more diverse types of GUI data (e.g. single parents, young parents, old parents, same-sex couples) than are currently provided in order to be able to tailor research to their specific remits, which could change rapidly across time. Stakeholders also mentioned that GUI data could provide a common starting point across organisations for articulating issues and policies of child welfare and well-being that clearly outline remits, responsibilities, and funding frames for evaluation.

All stakeholder groups (Government Departments, State agencies, NGOs, and HEIs) recognised the importance of GUI data for generating policy-relevant knowledge regarding child and family welfare and well-being.

# 5.5 Stakeholders' views of a new GUI birth cohort

## 5.5.1 GUI participant groups

*Summary:* Stakeholders recommended including participants from the medical profession (e.g. doctors, nurses, and midwives), pregnant mothers, members of diverse household structures (e.g. grandparents, same-sex couples, siblings, other children), teachers and school principals, different types of childcare providers, and children and parents from minority groups (e.g. special needs; cultural, social, and religious minorities).

The most proposed and discussed prospective participant groups in a new birth cohort fell into four categories: (1) health and pregnancy, (2) diverse families and households, (3) education and childcare, and (4) minority groups.

### 5.5.1.1 Health and pregnancy

Stakeholders discussed an interest in new participant groups from the healthcare sector, including GPs and hospital staff such as doctors, nurses, midwives, and experts working with children and infants with specific needs. Engaging with medical professionals in a general sense could provide an overall insight into the changing medical needs and challenges faced by children in Ireland.

Stakeholders also mentioned that it would be useful to collect information on pregnant mothers through prenatal appointment systems, parental support schemes, and training programmes. In addition, stakeholders suggested recruiting nurses and midwives in maternity hospitals as gatekeepers to introduce the GUI study to young or single mothers and other types of families. This could help recruit pregnant mothers from diverse backgrounds by getting them excited about the work and outputs of the GUI study.

> *"It would be useful to have the input of midwives early on. There's a lot of measurement in pregnancy with a whole range of information available at the clinical level which is valuable. Then you get information about [mothers'] education, their stressors, their conditions, everything about their work life, and their relationship information."* Stakeholder 6 (HEI)

### 5.5.1.2 Diverse families and households

Other suggested participant groups included diverse families and households like single parents, same-sex couples, younger parents, older parents, step-parents, parents with special needs, intergenerational households, and more.

> *"People sharing parenting, people step-parenting, and any other kind of diverse family situation. We need to understand more about the arrangements that people live in. What's that like from a child's perspective? We don't have a lot of knowledge and examples around that in Ireland."* Stakeholder 11 (NGO)

Stakeholders also suggested that a broader range of participants living in or sharing the same household as the child or infant could be involved in the GUI study; for example, grandparents, siblings, and other children.

Another point of interest was transitional family arrangements, which are a unique dynamic currently captured by the GUI study.

> *"Transitional family arrangements are really important for children. Sometimes partners of parents go in and out of a child's life. Who are the significant adults for a child? All these family relationships are fascinating."* Stakeholder 11 (NGO)

### 5.5.1.3 Education and childcare

Stakeholders indicated the importance of capturing data within the education sector by involving school principals and teachers.

> *"Then there's other groups that have significant interactions with the child. So, you're looking at teachers in the child's school and you're looking at principals in a child's school who are going to be quite influential in child outcomes, they're really important providers of insight about the kinds of things that the children are experiencing."* Stakeholder 14 (HEI)

Stakeholders also discussed the potential for further investigating GUI participants' experiences of diverse forms of childcare, including childminders, au pairs, nannies, and centre-based or in-home provision. Stakeholders differentiated between two main groups of possible caregivers: those working professionally (with a PPSN) as childminders, and those who do not work registered under a PPSN. The latter group includes undocumented migrant workers, as well as others who take care of children without contractual obligations (i.e. family members, neighbours, friends). This was framed within the context of the government's commitment to the *First 5* strategy and work around regulating childminders by 2028 by bringing them into a State-supported regulatory environment.

#### 5.5.1.4 Minority groups

Stakeholders stressed a need to include children and adults from minority groups within the general population as participants in a new GUI birth cohort study. Commonly mentioned minority groups were children and parents from the Traveller and Roma cultures and from other cultural minorities, including people of African origin or Muslim faith. In addition, people with specific social or developmental needs due to disability or lack of resources were also considered as belonging to minority groups.

Stakeholders recognised the challenges inherent to recruiting a nationally representative sample of people from cultural, social, and developmental minority groups.

> *"I'm not sure if it would be appropriate to use those [support] groups to access people [with disabilities] …there is a danger of missing the people who aren't in those groups…. [It would be helpful] if there was some way to be present in their community for an open talk, advertising the study, and creating a greater impact compared to sending an invitation to the door."* Stakeholder 18 (State agency)

Stakeholders suggested reaching out to cultural minority participants through existing organisations and networks that have established contact with those minority groups. Tusla was mentioned because of its statutory responsibilities and its interactions with individuals. Smaller groups such as grassroots movements, like Roots in Africa -Ireland and Black and Irish, were also mentioned as potential gateways to new GUI participants. Other networks connected with the Traveller community can be found in healthcare projects such as the All-Ireland Traveller Health Study: Our Geels.

Strategies for recruiting people from cultural minority groups were discussed with respect to the assumption that some prospective participants might lack trust in the

Irish government and/or researchers and may see little value in their participation due to a lack of tangible output of relevance to them.

> *"Because they've [GUI participants from cultural minorities] seen so little to link between the time and energy that people give to supporting research – whether as organisations or individuals – on one hand and any outcome for their cohort or people that they are working with, on the other hand."* Stakeholder 17 (NGO)

Another group mentioned by stakeholders was children in contact with non-universal service providers (i.e. the legal system), a family court, mediation, or another form of effective dispute resolution. As some of these children are receiving services under the remit of NGOs and State agencies like Tusla, it was suggested to further collect information through their existing networks.

> *"The only thing I could think of engaging them [children of minority groups] is presence. So, if you want to recruit from particular communities or particular groups, I would say having a presence in a community [is key]."* Stakeholder 15 (HEI)

Disability was recognised as a potential barrier to participating in the GUI study due to limited cognitive awareness, limited literacy, or physical disability making it difficult for individuals to fill out surveys. A suggested strategy to minimise this barrier was for the GUI study to adopt socially inclusive methods by having a socially inclusive research tool. Here, 'socially inclusive' refers to accessibility for all members of society including those with vision or hearing impairments, those with literacy difficulties, and neurologically diverse individuals.

### 5.5.2 GUI content areas

*Summary:* Stakeholders' suggestions of relevant topics for inclusion in a new GUI birth cohort survey were the long-term impact of the COVID-19 pandemic; housing; local environments; friendship networks; community social networks; online environments and behaviours; parenting; early childhood education and care; parental financial arrangements; children's health; racism and discrimination; social disadvantage and poverty; disabilities and special educational needs; and children of families participating in government supports and educational provision programmes.

### 5.5.2.1 Social, environmental, and digital contexts

Stakeholders called for information provided by a new GUI birth cohort on the long-term impact of the COVID-19 pandemic. COVID-19 has had a significant impact on Ireland's policy development, as many policies have been delayed primarily due to the importance of work related to COVID-19.

> *"We have issued a progress report on our latest policy advice, and that progress report was published almost two years ago now at this stage. The finalisation of that advice has been delayed with COVID and various other reasons, but [the policy is] probably one of the biggest pieces of policy advice we've ever published."* Stakeholder 18 (State agency)

Further information that stakeholders wanted to see in a new GUI birth cohort survey centred around housing type (i.e. subsidised, private rental, owner occupied), the building's structural qualities, the social structures of groups of people living in the houses (i.e. cultural or income diversity), and the impact of these issues on children's well-being.

Discussions around the built environment extended to children's access (or lack thereof) to safe and secure play areas, green spaces, and common spaces.

Children's peers and friendships were also mentioned as important content domains. Suggestions included gathering information on a maximum of six of each child's friends, their age, and the composition of the friendship group. Further suggestions were to study who children mixed with outside of school and the structures of these social networks. The role of the broader relational context was also suggested as a topic of interest.

> *"The world does not stop at the parents, the grandparents, and their peers. The peers are one thing, but there's the wider community as well that these young people are embedded in. Some studies have tried to take this more seriously. For example, [the Millennium Cohort Study] was heavily clustered…[this design] did give them the capacity to be able to look at the called 'super output areas' which allow for very detailed work on the local environments."* Stakeholder 14 (HEI)

Here, the stakeholder refers to 'super output areas', which are smaller areas that are clustered to improve reporting of small area statistics. While the GUI study included similar Small Area Population Statistics (SAPS) produced by the CSO in the Cohort '98 Researcher Microdata Files when the children were aged 9 years, SAPS have not been used since. Finally, more information about the online environments to which children are exposed was also called for. Stakeholders suggested using these online environments to directly collect data on participants' behaviours.

### 5.5.2.2 Parents and childcare

Parenting and relationships between parents and children were mentioned as extremely important topics for a new GUI birth cohort survey.

> *"The parents and child caregiving relationship, the quality of that. The attachment, relationship, normal attachments, development, that kind of stuff. And the psychological well-being of the parents, the primary and secondary caregiver. If there is a crisis going on in the family, how did the family manage that, and the impact that has on the parenting of the child and the child's responses to that [crisis-related parenting]."* Stakeholder 12 (State agency)

Another topic of interest was childcare and education in early childhood, due in part to the increased State support for access to preschools and other recognised early childcare settings.

> *"Things have changed in Ireland, which is another reason why we need a new birth cohort, there are now more children than ever in early learning and care settings because of the free preschool, those two free preschool years. Point being is that more children are now spending longer in those settings. You won't understand children's lives if you're missing three-quarters of their day."* Stakeholder 5 (NGO)

Finally, stakeholders discussed the policy relevance of studying parental financial arrangements in the cases where one parent is non-resident in the household.

> *"Our policy view would be that child maintenance should be treated as [a] non-means tested non-tax payment for the child, but in practice many families have either private arrangements or there isn't child maintenance being paid…it's almost a policy blind spot because it's dealt with privately and policy-makers are responding to it, so there's lots of scope for Growing Up in Ireland to have the same impact as it did around education and childcare going forward."* Stakeholder 11 (NGO)

### 5.5.2.3 Child-centred studies

Stakeholders discussed studying children's health as an outcome that is influenced by a range of social determinants such as accommodation, education, and experiences of racism and discrimination.

The impacts of poverty and social disadvantages were considered as important for further child-centred studies. Stakeholders suggested investigating the influence of geographical, institutional, and individual factors and their intersectionality to better understand how these can disrupt or drive poverty.

> *"In education we have DEIS [Delivering Equality of Opportunity in Schools] schools in certain areas that provide additional support. It would be very helpful for us to know about children who are in poverty who are not in DEIS schools, to understand what services they are receiving, or are offered."* Stakeholder 1 (Government Department)

Most participants suggested including specific COVID-19 measures to capture how children experienced the global pandemic and to evaluate potential impacts on their lives.

> *"The pandemic has become such a fundamental focus in all our minds and the well-being of children is absolutely critical in all areas of their life like their general experience, their friendships, their social lives, and their family lives."* Stakeholder 18 (State agency)

Furthermore, stakeholders drew attention to children's subjective perceptions and how these were central to children's lived experiences in homes, schools, and other settings.

> *"It's probably more like how [the children] relate to their other family members, other members of the household, what's the quality of that relationship they have with their parents or with whoever is living in the household."* Stakeholder 12 (State agency)

Children's experiences with racism in education and children's subjective perceptions and accounts of racism were also said to warrant more focus. One area of interest within this topic included ethnic minority children's perceptions of their own home and family structure. Stakeholders also discussed the importance of capturing the perceptions and experiences of children with disabilities who go to special schools and classes, and of children with disabilities who attend mainstream schools.

## 5.5.3 GUI methodology

*Summary:* Stakeholders also had suggestions regarding the use of social media to promote the GUI study, linking administrative data, building closer relationships to increase knowledge and value of the GUI study in specific cultural communities, and actively involving community members in GUI research.

### 5.5.3.1 Communicating results and impacts

Stakeholders suggested a multi-channel approach to successfully communicate with a public audience, including radio advertisements or interviews, funder and researcher appearances on television and news shows, and printed and digital news and media formats. Stakeholders from Government Departments also liked receiving their information in both a detailed and brief report form.

> *"Tweet it! Things get tweeted...academic journals speak to other academics and people's publishing careers and promotions and that's wonderful, but the information needs to be out there in a way that is digestible, and it needs to be up in the platforms that people access. Podcast! Include parents or kids that are involved in the GUI to share their experiences and how great the outcomes are."* Stakeholder 2 (Government Department)

For most stakeholders, effective communication of GUI findings hinged on fostering trust with the researched groups and communities.

> *"[The researchers] would be from the communities and groups [they wish to research] to begin with. The biggest challenge is creating conditions for people so they begin to see the value of engaging with the process. So, there's a bit about the sponsors of the research being able to put the cards on the table and being prepared to provide evidence on how this research could make a difference."* Stakeholder 17 (NGO)

Despite having discussed the importance of making children's perceptions and experiences central to the development of all questions, no stakeholder mentioned how the GUI findings should be communicated to children.

#### 5.5.3.2 Linking data sources for more insights

Data linkage came up again in the interviews when stakeholders began discussing methodological directions for a new GUI birth cohort. Stakeholders recognised the promises and challenges of linking together different administrative data (healthcare, education, and local level data) with GUI data sources.

> *"It is a lot of work to pin together information that is collected in ways that are structurally very different. So, if you're working with education or health management information systems across counties and they simply don't match, you have to do the hard work to not just extract the information from one source but to bring that together in a meaningful way."* Stakeholder 10 (HEI)

Healthcare systems that could contribute linked biomedical data were among the most important data systems that participants wished to see linked in a future GUI study and across other datasets and systems. This issue was raised by four stakeholders who were interested in the GUI collecting blood samples and blood types, tissue samples, and genotypes.

> *"We simply don't have any data on that in Growing Up in Ireland around that [genotype] and that's very important. No child studies should be without those measures [biomarker and genotype] because they're just too important in terms of looking at outcomes for the child and to measure processes as well. So a study that had those I think would be a real step forward for Ireland, if we could get hold of those data."* Stakeholder 14 (HEI)

Stakeholders also proposed that the CSO could manage these linked datasets, thereby matching school ID, neighbourhood ID, and socioeconomic clustering of participants' environments for more insights into children's environments.

#### 5.5.3.3 Inclusive, community-led research

The inclusion of community members in the design and execution of research was mentioned by nearly all stakeholders as being important for the GUI study. Stakeholders noted how community members could engage in collaboration with research professionals including the GUI study team, HEI employees, and PhD students.

> *"It's not just a question about, 'How do you drag [GUI participants] in?' It's a question of who owns the research? How is it articulated? Who's involved in*

> *this? How the questions are framed, and what the potential outcomes are for the groups. All of those questions need to be considered, not just by those leading on the study on their own but also in association, both with the organisations and the groups that are directly involved."* Stakeholder 17 (NGO)

Additionally, nearly all stakeholders considered community-oriented research as an opportunity to co-create and develop better research sample frames, methods, and tools which could better capture Ireland's dynamic demographic change over the past two decades and those to come. Consultation with special interest groups, community representatives, parents, and, of course, children and young people is an important aspect of the GUI study that should be preserved. These consultations should include discussions of target samples, appropriate questionnaire modes, and alternative data collection tools.

## 5.6 GUI data access

### 5.6.1 Education and training in secondary data analysis

Summary: Stakeholders suggested investing in more training and education for secondary data analysis in Ireland.

Stakeholders from all groups discussed the importance of increasing the number of secondary data analysts in Ireland through specific training programmes, such as online training sessions, institutional workshops, and webinars.

> *"We've got to have more courses on the key skills that people require – not just statistics but data analysis or data manipulation, data analytics. People need to have the courses available…cheap enough and subsidised enough that people can take them. They've got to be run regularly, they've got to be available all over the country, and they've got to be available on a regular basis, to make sure that people have an opportunity to take those courses when they need them."* Stakeholder 14 (HEI)

Higher education programmes for quantitative and secondary data analysis, such as specialised master's and PhD programmes, were also viewed as a potential solution for what six stakeholders identified as a lack of national expertise in analysing secondary data.

> *"I think there have to be PhD programmes embedded into colleges and widely used to build the expertise. We're possibly in the early stages in terms of having that expertise here in Ireland and we don't have a long tradition of longitudinal studies like our neighbours over in the UK."* Stakeholder 13 (Government Department)

### 5.6.2 Different modes of accessing GUI data

*Summary:* Stakeholders spoke positively about accessing GUI data using reports and data analysis archives. Stakeholders suggested that an online dashboard providing interactive data analysis would be useful.

Stakeholders discussed accessing future GUI data through three mechanisms. First, stakeholders were very positive about receiving short reports summarising data findings and interpretations of the results. Second, stakeholders who wanted to access the raw data suggested using a graphical user interface to better understand and handle large data tables in preparation for more specific statistical analyses. A given example of such graphical user interfaces (or dashboards) was the Higher Education Authority website. Third, stakeholders suggested having GUI staff members provide various services ranging from commissioned research to running an online matching platform for researchers and organisations. For example, in an online matching platform, researchers could list their interests and availability, organisations could list their analysis priorities and needs, and the two groups could browse and contact potential partners for collaborative analysis projects. Fourth, stakeholders also advocated for continued access to GUI data.

> *"This is a publicly funded dataset and, in my view, therefore, should be as widely available and [in] as accessible a manner as possible."* Stakeholder 16 (State agency)

### 5.6.3 Lack of organisational resources for data analysis

*Summary:* Stakeholders from NGOs described in detail their lack of staff and expertise in making use of GUI data in house. More collaboration between HEIs and organisational research teams, and more support for in-house research initiatives, were called for to make better use of the GUI data to inform practice.

Stakeholders discussed their organisational resources and access to datasets or professional staff for research and analytical tasks. Stakeholders from Government Departments and some State agencies were reasonably satisfied with their ability to access the GUI data for commissioned research, but expressed concerns about not having enough time to engage with GUI content in their day-to-day work or position.

> *"We might commission a piece of research from time to time – we're not a research institution – where we think there's a gap in information, or we need to understand more about a particular issue affecting the rights of children, or a particular cohort of children."* Stakeholder 16 (State agency)

In comparison, stakeholders from NGOs and other State agencies expressed a lack of capacity to use GUI data for their own research purposes.

> *"We can't afford it. We have one and a half researchers. We can't afford the resources Growing Up in Ireland has but we'd love to be able to utilise the research."* Stakeholder 3 (NGO)

All NGO stakeholders described having a lack of analytic capacity related to insufficient financial support for research staffing or funding for commissioned research. Oftentimes, these organisations faced the challenge of obtaining and managing resources to engage in their day-to-day activities while still engaging in political advocacy and research.

> *"It's difficult because being a national level organisation – and there's only a few in the area of migration – you're kind of asked to cover everything."*
> Stakeholder 7 (NGO)

NGOs and State agencies that conducted research reported their research as mainly being small-scale, non-representative, and qualitative. Related to this point, stakeholders indicated the importance of developing a future research labour force, ideally through cross-fertilisation with the higher education sector.

> *"It's like a jigsaw of knowledge and we need everyone. We need a lot of skills and we need especially creativity in the research process and integrity. So we really need a stream of young people coming along with new ideas and new approaches and so it must be multidisciplinary."* Stakeholder 16 (State agency)

## 5.7 Conclusion

To better understand how expert stakeholders in Government Departments, State agencies, NGOs, and HEIs perceive the value of a potential new GUI birth cohort, a stakeholder consultation was conducted. The consultation involved 18 stakeholders, 14 of whom were female, who were spread across Government Departments, State agencies, NGOs, and HEIs.

### 5.7.1 A new GUI birth cohort is needed

The importance of a new GUI birth cohort was the most salient point made by stakeholders, with their views on the current GUI studies being largely positive. A new birth cohort was perceived as extremely important due to the rapidly developing organisational remits and policy landscape.

### 5.7.2 GUI can be an integrative force for good

Stakeholders spoke about the value of the GUI study for establishing new directions for both Government Departments and NGOs to integrate policy and support services in key areas (e.g. refugee children).

### 5.7.3 Focus on diversity

Stakeholders emphasised the importance of capturing Ireland's social and cultural diversity in the GUI study. Their suggestions supported the inclusion of existing topics (e.g. parenting) and pointed to new directions for analysing child-centred data (e.g. children's perception of being in a government programme like social welfare). In addition, stakeholders reiterated the importance of including participants who reflect the whole spectrum of human diversity (e.g. ethnic minorities, people with additional educational needs and disabilities) and including them in sufficient numbers for statistical analysis.

### 5.7.4 Focus on context

Stakeholders suggested that a new GUI survey could cover more of the broader cultural context surrounding children's lives, including children's peer networks, broader community activities and stratification, and geographical and small area statistical data. This recommendation also pointed to the inclusion of new participant groups, including midwives, pregnant mothers, siblings, and diverse families. The recommendation also pointed to studying the dynamics of Ireland's changing population and the social interactions that affect children within their communities.

### 5.7.5 Integration of communities into GUI activities

Suggestions for participant recruitment included increasing connectivity with local communities and minority cultural groups. The idea of further community integration extended to involving local communities in GUI research design and analysis.

### 5.7.6 Data linkage

Stakeholders emphasised the importance of data linkage at national and sub-national levels for improving the analytical promise of GUI data, and they also proposed a need for change in Ireland's approaches to integrated national data systems.

### 5.7.7 Improved data analysis capacity

The stakeholders made a call for increasing analytic capacity nationally, with suggestions for new higher education programmes, more training in secondary data analysis, and interactive data analysis for the public (e.g. online interactive tools). Stakeholders from NGOs were most limited in using GUI data due to a lack of research staff and a lack of staff expertise. Suggestions were made for HEI and NGO partnerships to assist with data analysis and training.

> *"A new child cohort study gives us a fabulous opportunity to extend what we did with Growing Up in Ireland. Think about new dimensions and new methodologies! Technologies have moved on. So, let's look for new ways of doing things, and see if we can make this new GUI cohort study an innovative one that will break new ground internationally."* Stakeholder 14 (HEI)

# Appendices

## Appendix 5A

Stakeholder Consultation Interview Protocol

Questions

1. Please explain to me a little about your professional role in your organisation.
2. What is your professional experience relating to infants' and children's development and well-being?
3. How familiar are you with the Growing Up in Ireland study?
    - (a) (If familiar) can you give me any examples of how you have become familiar with the GUI study?
    - (b) (If not very familiar) read the GUI introduction brief.
4. What are the needs of your organisation for good-quality evidence on infant and child development and well-being?
    - (a) Why does your organisation have these needs?
    - (b) What type of evidence does your organisation most need?
    - (c) Why do you say this?
5. What are the needs of your organisation regarding policy on infant and child development and well-being?
    - (a) Why does your organisation have these needs?
    - (b) What type of policy does your organisation most need?
    - (c) Why do you say this?
6. Who are the key groups that the GUI should gather data on for a new infant cohort study?
    - (a) Can you explain why you think that X group should be included in the study?
7. What are the best ways of engaging these key groups in the GUI study?
    - (a) Can you give me some reasons for why you think these ways would be most effective?
    - (b) What are some potential barriers to participation for these key groups?
8. What are your views on the key content areas that a new GUI infant cohort study could explore?

   (a) What are some specific topics you would like to see included, bearing in mind of your professional role?
   (b) In what ways is that topic relevant to your organisation's evidence or policy needs?
   (c) Why do you think that topic should be included in a new GUI infant cohort study?
9. Have you ever accessed or analysed the data from the current GUI cohorts?
   (a) If so, why did you access or analyse these data?
   (b) If not, why not?
10. Can you think of anything that might facilitate better access and use of the GUI data for organisations or sectors like yours?
11. Can you think of any specific challenges that organisations or sectors like yours might face in accessing and analysing the GUI data?
12. Do you have any final comments or observations you would like to add before we finish this interview?

# Chapter 6: Options and conclusions

## 6.1 Chapter overview

In response to DCEDIY's call for a scoping review to inform a potential new GUI birth cohort, this chapter presents a synthesis of the main options and best practices uncovered by our review of the academic literature and named birth cohort studies, and from the stakeholder consultation. This chapter is organised to give up-to-date coverage of the key elements core to the development, design, and implementation of birth cohort studies: study objectives (Section 6.2); conceptual frameworks (Section 6.3); data collection methods (Section 6.4); sampling and recruitment (Section 6.5); participants and respondents (Section 6.6); data collection domains and measures (Section 6.7); and data analysis and outputs (Section 6.8). In each of these sections, we conclude with a table of options for the Irish context which details the available options, their benefits, and their disadvantages. The chapter concludes with an overview of why a new GUI birth cohort study is needed.

### 6.1.1 Key lessons

#### 6.1.1.1 Study objectives

Ireland's ever-changing social and cultural diversity, together with relevant policies and services, should be taken into consideration when defining the objectives of a potential new birth cohort study.

#### 6.1.1.2 Conceptual frameworks

Conceptual frameworks give context to birth cohort data and guide data collection and analysis. Selecting a conceptual framework for a potential new birth cohort should consider the capacity of the framework to grow with the study and adapt to different life transitions (i.e. from childhood to adolescence).

#### 6.1.1.3 Data collection methods

Over a decade since GUI was first launched, technological innovations offer a broad menu of considerations for the development of a new GUI birth cohort. New technologies (e.g. web-based surveys, smartphone apps, fitness trackers) could facilitate the remote collection of birth cohort data at scale. However, implementing new technologies is often met with new challenges such as lower response rates and complex ethical, security and data protection considerations.

### 6.1.1.4 Sampling and recruitment

The sampling frame from which the population is drawn influences the overall sample quality. Potential sampling frames that might be considered in the development of a future birth cohort in Ireland include the Child Benefit Register, the Birth Notification Form (Form BNF/01), and the ongoing Maternal and Newborn Clinical Management System (MN-CMS). The potential use of these administrative records is contingent on factors such as public acceptability, GDPR compliance, and administrative feasibility.

### 6.1.1.5 Participants and respondents

As echoed across the stakeholder consultation, the robustness of the data produced by birth cohorts is dependent on the ongoing representativeness of the population and the inclusion of traditionally hard-to-reach groups. A targeted sampling strategy (e.g. oversampling, stratification by geographical area) is needed in order to ensure the representation of members from policy-relevant subpopulations.

### 6.1.1.6 Data collection domains and measures

The possibility of measurement harmonisation should be given consideration when identifying measures to cover target domains. For instance, cross-cohort harmonisation with previous GUI studies would provide comparable data to shed light on time trends and history effects, as balanced with the need to capture social change and increased diversity and to incorporate new technologies and data collection instruments. Cross-study harmonisation with international studies (e.g. the Early Life Cohort Feasibility Study (ELC-FS), Growing Up In Digital Europe (GUIDE)) is another option that would facilitate cross-cultural comparisons with other countries.

### 6.1.1.7 Data analysis and outputs

The development of a new birth cohort will entail novel data analysis strategies and new approaches to data collection and recruitment, together with greater sample diversity. With these advances in mind, training in terms of how to access and utilise the data is imperative from the first wave of new data outputs.

## 6.2 Cohort study objectives

*Summary:* Given Ireland's ever-evolving social and cultural diversity, a potential new birth cohort study should strive to inform relevant policy and support services and guide the decisions of government and non-government organisations.

The range of objectives across the named birth cohort case studies in this review were varied, with useful learnings for potential future GUI study objectives. One MCS objective was to collect information on previously neglected topics, such as the role of fathers, non-parental childcare, and ethnicity (Joshi & Fitzsimons, 2016). Another MCS objective was to provide evidence for use of the national evaluations of Sure Start and the Children's Fund (Joshi & Fitzsimons, 2016).

These points are also connected to insights offered by stakeholders. A new birth cohort study was considered necessary due to the rapidly developing organisational remits and policy landscape. Some participants also addressed the value of the GUI study for helping to inform policy and support services in key areas (e.g. child protection, adoption, family integration, youth services, refugee support services, equality policies). The GUI study was thought to have continued potential to be an integrative force for good in establishing new directions for government and non-governmental organisations. Stakeholders also addressed the importance of capturing Ireland's social and cultural diversity in a future GUI study, with a continued focus on the inclusion of various minority groups in a new survey design (e.g. ethnic minorities, individuals with special educational needs and disabilities). They also recommended collecting data on the broader cultural context of children's lives. For example, consulting medical professionals and including new participant groups (e.g. pregnant mothers, siblings, members of diverse family structures), and the study of dynamic processes and social interactions that impacted on children in their communities (e.g. personal, family, situational and environmental stressors) (Grosser *et al.*, 2016).

**Table 6.1** Options for the Irish context: Study objectives

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Collect information that can be compared to earlier GUI cohorts, in order to identify time trends. | Using ex-ante harmonisation of measures to collect the same data as previous GUI birth cohorts (see Chapter 1 for an overview of existing GUI study objectives). | Comparability of measures across all GUI cohorts. | Reduced scope for including new measures. |
| Collect data on the broader cultural context of children's lives. | New data on broader contexts could be gathered by adding new participant groups and studying dynamic processes and social interactions that impact on children within their communities. | Adding new participants (e.g. siblings and extended family members, healthcare professionals) would produce a fuller picture of children's lives. | This would add complexity and cost to data collection. |
| Capture Ireland's social and cultural diversity. | A future GUI study could achieve this by adopting a continued focus on the inclusion of minority groups in a new survey design (e.g. ethnic minorities, people with special educational needs and disabilities). | Oversampling specific minority groups produces large enough samples for within-group analyses. | Since ethnic minority groups are diverse, with different composition and experiences (e.g. Black, Asian and minority ethnic groups, Travellers, high-skilled migrant workers, immigrants from developing countries, and refugees from Ukraine and other international wars/crises), it will be challenging to identify and oversample these different groups. Different categories of ethnic minority groups should be clearly defined in advance when determining the target sample. In the case of cells with a small number of individuals from a particular ethnic minority category, further intervention is needed to increase the cell size and avoid insufficiently sized subgroups of interest for analysis. |

## 6.3 Cohort study conceptual frameworks

*Summary:* Conceptual frameworks should be capable of growing with the study and/or adapting to different life transitions (i.e. the transition from childhood into adulthood).

The conceptual or theoretical framework is the soul of every research project. It determines how a given researcher formulates their research problem, how they go about investigating the problem, and what meaning they attach to the data accruing from such an investigation (Imenda, 2014, p. 185).

A wide array of conceptual frameworks are used across birth cohort studies, and these generally take a social science or a health perspective (Wang *et al.*, 2021). However, as identified in this review, there is clear variation in the extent to which study documentation makes the conceptual frameworks explicit. For example, with regard to the DNBC study, traces of the conceptual framework are mentioned in the documentation; however, the conceptual framework is not explicitly described. Conversely, the GUI study has a well-defined and clearly explained conceptual framework which provides context to the data and guides data collection and analysis. One important consideration, particularly for longitudinal birth cohort studies, is that the conceptual framework can grow with the study or can be adapted when there are life transitions (i.e. the transition from childhood into adulthood). As Wang *et al.* (2021) identified in their work to develop a conceptual framework for the GenV study, many frameworks are not mutually exclusive. Instead, in complex omnibus studies which measure social science, health, and economic outcomes, multiple frameworks may be required to complement and contribute to the overall evaluation of the study goals.

**Table 6.2** Options for the Irish context: conceptual frameworks

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Build upon the existing GUI conceptual framework. | The GUI study is underpinned by Bronfenbrenner's bioecological model, which is adaptable to new contexts and life course stages. | Consistency with the existing GUI study.<br>Potential to adapt Bronfenbrenner's model to incorporate new research topics and participant groups. | Forgoing an opportunity to develop a new life course framework specifically for the new study. The development of a new, purpose-built framework could have a significant effect on focusing the direction, and ensuring the overall merit, of a new GUI cohort. |
| Develop a new framework that will consider further life span and intergenerational elements in the study in order to align the conceptual framework with the new study objectives. | Following the example of GenV, identify specific criteria when selecting from a range of existing life course frameworks to inform a customised framework for the new study.<br>The assessment criteria used by Wang *et al.* (2021) in selection of GenV's primary framework included:<br>• **Scope/hypothesis:** broad and integrated<br>• **Broad and multidimensional:** not focused on narrow environmental exposures<br>• **Outcomes:** a consideration of the wide-ranging aspects of physical, mental, and social well-being and not simply the absence of disease<br>• **Life course perspective:** acknowledgment that outcomes are constantly embedded and influenced by time and place across the life span and generations<br>• **Mechanisms/pathways and interactions:** a consideration of how health and disease outcomes develop from adaptive, multilevel, and reciprocal interactions between people and their environments<br>• **Multi-age:** the framework is applicable to all children and adults | A transparent and systematic process for developing a framework specifically for the new study. | Additional effort involved in purposefully developing a new framework. |

## 6.4 Cohort study data collection methods

Data collection methods have evolved dramatically in recent years. This section will explore the data collection options available, and their associated advantages and disadvantages, through a discussion of CAPI and CASI, web-based surveys, smartphone apps, wearable fitness monitors, biomedical data collection and administrative data linkage.

### 6.4.1 CAPI and CASI

To date, the GUI study has relied predominantly on CAPI and CASI techniques. Face-to-face interviews have been conducted with cohort members and their families via CAPI, whereas questions pertaining to more sensitive topics have been included in a CASI self-completion module on the interviewer's laptop. Given the COVID-19 pandemic, online data collection was used with Cohort '08 and Cohort '98 in December 2020. At present, the current 25-year-old Cohort '98 pilot study will explore the feasibility of online survey/telephone interview modes, compared to traditional CAPI/CASI (Smyth *et al.*, 2022).

**Table 6.3** Options for the Irish context: CAPI and CASI

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| CAPI | The interviewer uses an electronic device (e.g. iPad, laptop) to record the participants' responses to interview questions. | • Used effectively in previous GUI waves.<br>• Enables the routing of questions (e.g. skipping non-applicable questions).<br>• Facilitates the use of logic checks, skip patterns, and validations (Thornton *et al.*, 2013). | • Requires face-to-face home visit.<br>• More costly than web-based surveys.<br>• Requires electricity and stable Internet connection. |
| CASI | Participants complete confidential questionnaires using an electronic device. | • Increases the efficiency and accuracy of data collection as participants may be more inclined to disclose sensitive information using this method which provides increased privacy.<br>• Automatic branching reduces participant burden and ensures that the participant navigates the questionnaire in the intended manner.<br>• Range rules and consistency checks reduce the likelihood of rogue answers.<br>• CASI increases the perception that data remain confidential (Thornton *et al.*, 2013). | • Participants may still be unwilling to disclose sensitive information.<br>• Requires electricity and stable Internet connection.<br>• Relies on participants' familiarity with computers and electronic devices. |

## 6.4.2 Web-based questionnaires

*Summary:* Web-based surveys are efficient and yield considerable cost savings but tend to achieve lower response rates than traditional survey modes.

Web-based questionnaires are significantly cheaper and more efficient to administer when compared with traditional face-to-face interviewing, CAPI, and CATIs (Cocchi, Giovinazzi & Lynnet, 2019; Kelfve *et al.*, 2020). Nevertheless, web-based questionnaires tend to achieve lower coverage and response rates compared to other survey modes (Cocchi, Giovinazzi & Lynnet, 2019). It has been suggested that the abundance of online surveys in circulation nowadays has served to increase respondent burden and fatigue, in turn decreasing response rates (Harrison *et al.*, 2020). Therefore, while web-based surveys are more efficient and yield considerable cost savings, their potential downside includes increased attrition and non-response (Cocchi *et al.*, 2019). To combat these potential lower response rates, recruitment methods (e.g. offering small incentives, regular contact with families between waves) and sampling approaches (e.g. boosted samples) should be considered to overcome potential sample bias due to attrition and non-response.

With regard to the Irish context, GUI participants in both the Cohort ’08 and Cohort ’98 studies were invited to participate in a special COVID-19 online survey during December 2020 (GUI, 2021). There was a 45% response rate among primary caregivers, a 38% response rate among Cohort ’08 participants, and a 33% response rate among Cohort ’98 participants. While overall response rates were significantly impacted by the online-only mode, the underlying trends in response patterns were comparable to earlier in-home interviews; for example, lower response rates among those from disadvantaged backgrounds (lower income, lower parental education, and social class) and among the older cohort of participants, both males and females, in Cohort ’98 participants (Kelly *et al.*, 2021). Notwithstanding a significantly lower response rate than for earlier waves, the web-based design did manage to achieve its primary goal of facilitating data collection that was convenient and meaningful within a brief period during which in-person data collection was not possible. Factual data were collected regarding the impact of the COVID-19 pandemic on employment, income, education and physical activity, alongside subjective indicators of stress and mental well-being (Kelly *et al.*, 2021). For this purpose, the use of web-based surveys was advantageous to the GUI team in providing fast access to relevant data, and empowering respondents to complete the survey remotely and at their own convenience through a variety of platforms, such as computers or mobile devices (Kelfve *et al.*, 2020).

**Table 6.4** Options for the Irish context: web-based questionnaires

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Web-based questionnaires | Web-based questionnaires have important implications for use within a new longitudinal birth cohort study, which should strive to maximise participation and ensure the quality and longitudinal integrity of gathered data, while simultaneously reducing costs and alleviating participant burden (Goodman *et al.*, 2020).<br>The optimal choice of survey mode (interview, paper, post, telephone, web, or a mixed mode) for a future GUI study may involve a trade-off between priorities and will depend on the objective of the new birth cohort study, alongside the target population, the representativeness requirements and the financial resources available (Kelfve *et al.*, 2020). | • Can be used as an alternative or supplement to other survey modes or as a stand-alone measure.<br>• Is efficient and would yield considerable cost savings.<br>• Facilitates remote data collection.<br>• Does not require an interviewer to facilitate data collection.<br>• Allows for a larger sampling frame for a broader target population. | • Achieves lower response rates than traditional survey modes.<br>• Could lead to higher rates of attrition and item non-response.<br>• Could lead to lower response rates among traditionally hard-to-reach groups.<br>• Could lead to greater respondent burden and survey fatigue due to the vast number of online surveys in circulation nowadays. |

### 6.4.3 Smartphone apps

*Summary:* The growing demand for mobile health (mHealth) apps presents researchers with opportunities to gather participant-generated data using smartphone apps.

mHealth apps have become hugely popular since the early 2010s, particularly among expectant and new mothers, among whom app use is exceptionally high, predominantly as a means of obtaining pregnancy-related information outside of clinical hours. There is a growing demand for apps that provide information relating to pregnancy and childcare, especially among first-time parents, with an extensive availability of apps that aim to educate and support women throughout their pregnancy and during the postpartum period (Cawley et al., 2020). Previous studies have also demonstrated the willingness of mothers to share information about their pregnancy via apps for research and service provision purposes (e.g. Deave et al., 2019; Radin et al., 2018). Relatedly, the new ELC-FS in the UK plans to test the feasibility of smartphone-based data capture by using the app BabySteps, developed at the University of Iowa (Cramer Development Incorporated, 2022), to assess developmental trajectories (e.g. infant sleep habits) and family processes (e.g. ecological momentary assessments (EMAs) of parental mood states) (CLS, 2021b).

Notwithstanding these research opportunities, the immense volume of personal data that can be collected using mobile technologies raises ethical, security and data protection considerations that should be of paramount importance to the usage of smartphone-based data capture in a future birth cohort study (Stone & Skinner, 2017). Moreover, the actual acceptability of different forms of smartphone data presents potential challenges for participant engagement (Ashing et al., 2018). Wenz, Jäckle & Couper (2019) observed that, among members of the UK Household Longitudinal Study's Innovation Panel, willingness to participate varies substantially depending on the type of mobile data collection activity involved. According to Wenz and colleagues (2019), this variation was associated with the characteristics of the different activities; fewer panel members would be willing to share the GPS position of their smartphone than they would be to take a photo for a survey or to complete a survey in a mobile browser. Similarly, more than one-half of the panel members stated that they would not be willing to download an app that gathers anonymous data regarding their mobile device usage. These findings indicate a greater overall willingness to take part in active data collection measures than to permit the passive recording of smartphone data (Wenz, Jäckle & Couper, 2019).

Furthermore, as technology is ever-evolving, the accuracy of smartphone-generated data must be thoroughly examined prior to a national roll-out of smartphone-based data collection. Important considerations include evaluating the accuracy of screen time data (e.g. an app being actively used versus simply being left open) together with considering potential variations across different smartphone models and operating systems (e.g. Android versus iOS).

**Table 6.5** Options for the Irish context: smartphone apps

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Smartphone apps | Drawing on widespread mHealth app usage, there are opportunities to transform birth cohort research within the Irish context. | • Smartphone usage in Ireland is at 95% for those aged 16–29 years, and 96% for those aged 30–34 years (Gibney & McCarthy, 2020).<br>• New parents are willingly downloading parenting and mHealth apps. With participants' informed consent, researchers could explore the feasibility of smartphone-based data collection embedded within popular mHealth apps.<br>• Smartphone apps can gather participant-generated smartphone data in a way that is less burdensome to participants.<br>• Smartphone apps can promote better long-term engagement with the study, as data can be gathered remotely, reducing or eliminating the need to attend a formal clinical setting.<br>• There is a possibility for data linkage through existing apps. | • The acceptability of smartphone-based research will depend on the type of data being collected.<br>• Participants are more likely to opt for active data collection measures (e.g. surveys, taking photos/videos) than the passive recording of smartphone data (e.g. app usage, browser search history) (see Wenz, Jäckle & Couper, 2019).<br>• There are complex ethical, security and data protection considerations.<br>• A significant investment would be required if developing a new, study-specific app.<br>• The accuracy of smartphone-based data collection may vary across different models and operating systems. |

## 6.4.4 Wearable activity monitors

*Summary:* Wearable activity monitors provide a more objective measure of physical activity than self-report measures.

Physical activity is vital to the health and well-being of children. The Irish primary school curriculum acknowledges this by recommending 1 hour of physical education each week, and schools supplement this with further exercise opportunities given during school break times and in extracurricular activities organised by the school (Growing Up in Ireland Study Team, 2018). Wearable activity monitors are increasingly being used to deliver a convenient and objective surrogate index for physical activity, validated by evidence-based steps per day categorisation along a physical inactivity/activity continuum (Cai, Tan & Ang, 2019). Accelerometers in particular are quite straightforward for participants to use and have been successfully implemented in previous birth cohort studies (Western *et al.*, 2014). Besides traditional research-grade accelerometers, technological advances have resulted in a wider variety of commercially available personal fitness trackers (e.g. Fitbit) and smartphone apps that can be utilised to record physical activity and sedentary behaviour. The sleeker appearance of Fitbits and smartphone apps could encourage participant compliance while continuing to gather valid data. Furthermore, these devices are more cost-efficient and reduce staff resource requirements associated with the distribution of traditional research-grade devices (Gilbert *et al.*, 2017).

**Table 6.6** Options for the Irish context: wearable activity monitors

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Wearable fitness trackers | The use of accelerometers could be a reasonably inexpensive means of collecting objective, policy-relevant data to inform interventions aimed at encouraging physical activity and attaining the associated physical and mental health benefits (Western *et al.*, 2014). | • Eliminates reliance on self-report data.<br>• Avoids recall and social desirability biases regarding physical activity.<br>• Personal fitness trackers and built-in smartphone accelerometers are now cheaper and more stylish than traditional research-grade accelerometers. | • Difficult to ensure participant compliance.<br>• Participants could share the fitness monitor device with another person.<br>• Could result in reactivity, as participants modify their usual exercise habits in direct response to their awareness of being observed. |

### 6.4.5 Biological data

*Summary:* International best practice points to the value of integrating genetic, epigenetic, and environmental variables with longitudinal survey data in order to promote aetiological knowledge of various childhood diseases (Giesbertz, Bredenoord & van Delden, 2016).

Many large-scale birth cohort study designs have constructed biobank repositories of biological samples to contribute to the study of the interplay between genetic, lifestyle and environmental factors in health and well-being across the life course (Townsend *et al.*, 2016). While biobanking has been previously carried out at a regional level in Ireland (e.g. the Cork BASELINE birth cohort study), the pressing need to include biomarkers at a national level in a future GUI study emerged across the stakeholder consultation interviews. The long-term storage of biological data also serves to future-proof a new birth cohort study by paving the way for future testing where analytic methods do not yet exist, or where sufficient samples are not yet available to make strong conclusions (Western *et al.*, 2014).

Notwithstanding the epidemiological value of collecting and storing biological data, paediatric biobank research is met with specific ethical issues. At inclusion, children often cannot provide (or are not legally capable of providing) informed consent for themselves, thus parental consent is required (Giesbertz, Bredenoord & van Delden, 2016).Given that biological samples will likely remain stored and used when cohort members become autonomous adults, a future birth cohort study must consider the scope of parental consent and the right of the child to assent or dissent to biobanking. Under the GDPR, consent must be obtained from the holder of parental responsibility for the child;[53] this consent must be explicit and 'opt-in' (DPC, 2019, 2021). Regarding re-consent, international practice recommends developing a clear protocol of obtaining the cohort member's own informed consent and providing them with the option to withdraw their samples once they reach the legal age of maturity (Giesbertz, Bredenoord & van Delden, 2016; Hens, Lévesque & Dierickx, 2011). While there remains substantial diversity in opinion, a growing consensus among the bioethics community indicates that, under certain circumstances, it would

---

[53] Defined as children aged under 18 years according to the Data Protection Act 2018, with 16 years being considered the "digital age of consent" in Ireland (DPC, 2019).

be ethical to disclose clinically significant findings if they could be of value to the cohort member (Kaye *et al.*, 2014). A good practice would be the development of a clear informed consent procedure whereby participants can also make an advanced decision concerning the return of incidental findings. Such information should be adapted to the type of biobank in question, and the consent process should involve a clear option to opt out, thereby providing participants with the choice to never receive information on incidental findings if they so wish (Aarts, Bunnik & Boeckhout, 2017).

**Table 6.7** Options for the Irish context: biological data

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Biological data | Collecting and storing biological samples at a national level could advance understanding of the social and genetic bases of diseases and has scientific and policy relevance, as Irish children are studied from infancy and throughout the life course. | • Certain biological measures (e.g. saliva) can now be obtained at home using straightforward, non-invasive procedures.<br>• Biobanking can facilitate current and future investigations into the role of genetic and environmental influences on human development across the life span. | • Given the legal and ethical issues associated with paediatric biobanking activities, a clear informed consent process would be needed to outline how the study plans to ensure safe storage of biological data, disclose incidental findings, and address issues of assent and re-consent.<br>• The collection of certain forms of biological data (e.g. blood) from young children would require specialist equipment and professional techniques.<br>• Partnerships with hospitals and other medical institutions, or home visits from medical professionals, may need to be contracted by the research team. |

## 6.4.6 Data linkage

*Summary:* Linkage of birth cohort and administrative data is used to drive policy and practice in many large-scale studies worldwide.

Linking survey data with administrative records (whether maintained by government or private entities) is increasingly recognised as an invaluable means of ensuring speedier data collection, reducing costs, alleviating participant burden, and enhancing the overall accuracy and completeness of longitudinal data (Jäckle *et al.*, 2018). Administrative records typically contain elaborate information (e.g. dates of hospital admissions, exam results, financial data), which would be burdensome for participants to recall and report accurately. Furthermore, data linkage is particularly advantageous for filling in gaps between waves or before the baseline wave and can facilitate ongoing data collection if cohort members are lost to follow-up for the purpose of non-response adjustments, where permissible under the GDPR and data protection legislation (Peycheva, Ploubidis & Calderwood, 2021).

Although linkage of administrative and birth cohort data is increasingly performed internationally, there is a limited centralised data infrastructure in Ireland (Walsh, Walsh, Mac Domhnaill & Mohan, 2021). Instead, isolated silos of information and a lack of cross-departmental collaboration prevents the secure, efficient transfer of administrative data (Gilbert *et al.*, 2018; HIQA, 2021). Findings from the stakeholder consultation emphasised the need for greater integration across different data management systems (e.g. health, education, social welfare) to enrich GUI data and provide a more holistic insight into the health and well-being of children in contemporary Ireland. While the culture of data sharing in Ireland remains quite closed, there has been a recent increase in the availability of data at an aggregated level (e.g. via the CSO StatBank and data.gov.ie) (Hanafin, 2020). The GUI study is carried out under an arrangement based on Section 11 of the Statistics Act, 1993, whereby the CSO is empowered to work with other public bodies for the collection, compilation, extraction, or dissemination of information for statistical purposes (CSO, 2022a). Accordingly, in line with the CSO's leadership role of the Irish Statistical System (ISS), one potential option for a future birth cohort study could be the development of a 'pathfinder' project in partnership with other public sector bodies (CSO, 2022b; Hanafin, 2020). Relevant to a future GUI birth cohort, a pathfinder project could be conducted wherein individual-level data from consenting cohort members are integrated with administrative data from a broad range of public sector bodies concerning the individual participants (e.g. the Revenue Commissioners, Department of Social Protection, the Higher Education Authority) and other issues of relevance (e.g. characteristics of their local area such as deprivation/affluence, population density) (Hanafin, 2020).

**Table 6.8** Options for the Irish context: data linkage

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Pathfinder projects | Pathfinder projects involve linking secondary datasets hosted by Government Departments with existing administrative data held by the CSO in order to yield aggregated analysis and outputs. | • Developed as part of the CSO's leadership role of the Irish Statistical System (ISS).<br>• Linked datasets can be established and managed by the CSO.<br>• Data linkage alleviates participant burden.<br>• Data linkage enhances accuracy and completeness of longitudinal data.<br>• Participants can be given options for linking or not linking data (this does not preclude their participation in the GUI study). | • May be complicated under the GDPR and Irish data protection regulations.<br>• Disparities across administrative datasets could limit their potential for statistical use.<br>• Common problems related to linking survey data to administrative data (non-consent; non-coverage; i.e. no data for a respondent in administrative datasets).<br>• Participants may not want to consent to data linkage. |

## 6.5 Sampling and recruitment considerations

This section will describe sampling and recruitment options for a future birth cohort study through an exploration of the sampling design and planning process, and the selection of a sampling frame.

### 6.5.1 Sampling design

*Summary:* Outlining a clear-cut sampling strategy ahead of time helps ensure the success of a future birth cohort study.

Specifying an appropriate sampling design in advance of beginning a study is imperative; the absence of a rigorous approach to sampling runs the risk of yielding samples that do not adequately represent the population at large (Martínez-Mesa *et al.*, 2016). A well-designed birth cohort study requires a substantial level of piloting, validating and strategic planning prior to going to the field (Doyle & Golding, 2009). Unfortunately, this is an expensive process and funding bodies are often reluctant to finance this critical preparatory work, as they would prefer for their investment to directly produce tangible results. Doyle and Golding (2009) recommend two potential options: (1) packaging the preparatory work as separate methodological projects with distinct end points and obtaining funding from one or more sources, or (2) the project receiving funding from bodies that fully acknowledge the significance of preparatory work (Doyle & Golding, 2009). For instance, the Innovation Panel of the UK Household Longitudinal Study exhibits the lasting benefits of investing in preparatory work. The Innovation Panel is a sample of 1,500 households, separate from the main study, used as a test-bed for innovative data collection methods and design features to ensure the ongoing success of the UK Household Longitudinal Study (Understanding Society, 2022). Methodological experiments are conducted to examine the effect of innovations in longitudinal survey methods by incorporating into their design experimental variation between different groups of panel members (Hanson, Westwood & Carpenter, 2018); for example, to examine response bias. Analysing interview data from these different groups facilitates assessment of the impact and relative merit of the different methodological approaches (Cernat & Sakshaug, 2021; Hanson, Westwood & Carpenter, 2018; Jäckle *et al.*, 2021).

Regarding a new birth cohort study in Ireland, the identification of a feasible sampling strategy could encompass a core component of preparatory work. One potential option is the establishment of an innovation panel to test the impact of different sampling frames on response rates and patterns. Alternatively, DCEDIY could commission individuals with methodological expertise in longitudinal research to explore potential sampling design options that would be feasible within the Irish context. This approach has been adopted in the UK. For example, a recent CLS paper offers an in-depth analysis of theoretical sampling design options that might be adopted by a new birth cohort study in the UK (see Sullivan, Joshi & Williams, 2020).

**Table 6.9** Options for the Irish context: preparatory work to identify a sampling strategy

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Engage in preparatory work to identify a clear sampling strategy. | Specify a feasible sampling strategy as a core component of preparatory work. | • Ensures a robust design that will yield a representative sample.<br>• Determines the value and feasibility of different methodological approaches. | • Very expensive and time-consuming to commission individuals to evaluate potential sampling frame options.<br>• Funding bodies are often unwilling to finance these discrete research projects. |

### 6.5.2 Sampling frame

*Summary:* The quality of the sampling frame influences the quality of the sample.

The chosen sampling frame from which the target population will be drawn has implications for the overall quality of the sample (Mohadjer *et al.*, 2016). Different countries vary on the availability and accessibility of suitable sampling frames, such as administrative records (e.g. healthcare records, electoral registers, or population records), address lists or area-based frames (Watson & Lynn, 2021).

Administrative records that might be used to select a sample within the Irish context include the Child Benefit Register, the Birth Notification Form (Form BNF/01) or the MN-CMS (once it has been implemented nationally) (see Chapter 3, Section 3.2.9 for further details). Yet, like all sampling frames, the actual applicability of these administrative records would be dependent on a multitude of factors, such as public acceptability and adherence to the GDPR and to data protection legislation in general, alongside the actual administrative feasibility of these records for sampling purposes (Sullivan, Joshi & Williams, 2020).

**Table 6.10** Options for the Irish context: sampling frame

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Child Benefit Register | A monthly payment to the primary caregiver of children aged under 16 years living in the Republic of Ireland. | • Financially advantageous; very few parents in Ireland do not claim Child Benefit.<br>• Universal payment: all parents in Ireland are eligible; income level and Pay Related Social Insurance (PRSI) payments do not influence eligibility.<br>• Each member of the target population appears just once on this centralised register.<br>• Contact information and social security number are contained in this register.<br>• Available in an easily accessible, electronic format.<br>• Previously successful in providing the sampling frame from which the GUI Cohort ’08 was drawn. | • The sampling procedure used to recruit the GUI Cohort ’08 predates the GDPR and other general data protection regulations.<br>• Accessing sociodemographic data and contact details from the Child Benefit Register might no longer be permitted today.<br>• Legal advice is therefore needed in order to determine whether this register could still be used as a potential sampling frame under modern data protection laws. |

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Birth Notification Form (Form BNF/01) | Completed with the parent(s) by hospital staff (for hospital births), or by a doctor or midwife (for home births), to ensure that accurate information is recorded. Contains the information to be recorded on the child's birth certificate (once registered within 3 months of the birth). | • Information recorded includes the infant's surname and forename; the time, date and place of birth; and the sex of the infant, together with details about the mother and father (forename, surname, date of birth, home address, PPSN, occupation, nationality, civil status).<br>• Completed for all infants born in Ireland (nationally representative sampling frame).<br>• Contains sufficient detail to facilitate initial contact with potential participants (e.g. home address) and subsequent tracing between waves (e.g. PPSN). | • Gaining access to these secure records would require consent from the HSE Civil Registration Service and/or individual maternity units/patients.<br>• Accessing these data may not be possible under the GDPR. |
| The MN-CMS | The MN-CMS refers to the ongoing implementation of a new electronic health record (EHR) for all women and infants being cared for in maternity, newborn, and gynaecology services in Ireland. | • Easy to access, single record on mother and infant that is electronically accessible across maternity hospitals.<br>• The MN-CMS has been successfully implemented at four maternity units in Ireland. | • Full roll-out of this system is ongoing, and may not be implemented for another few years.<br>• In its current phase, this system would limit the sample to only 4 out of the 19 maternity units/hospitals in Ireland – i.e. would not be nationally representative.<br>• Obtaining access to this system for research purposes might not comply with the GDPR.<br>• Prior consent would be required to access the contact details of patients for sampling and recruitment purposes. |

## 6.6 Participants and respondents

This section will outline options associated with the selection of a target population within the Irish context, alongside options for how a future GUI study could ensure the representativeness of the respondents targeted.

### 6.6.1 Prenatal sampling

*Summary:* Prenatal sampling could shed light on the critical influence of in utero exposures, but is met with inherent challenges within the Irish context.

A core consideration involved in the sampling design process is the clear identification of the target population of interest. Across European birth cohorts, an emerging trend is the commencement of sampling during the prenatal period, rather than recruiting cohort members at birth or in the months following childbirth (Pansieri *et al.*, 2020). From an epidemiological standpoint, sampling during pregnancy is critical to uncovering the influence of the early stages of development on life-long health and social trajectories (Charles *et al.*, 2020). At present, a potential sampling strategy would likely involve engagement with the Maternity and Infant Care Scheme, a national programme that delivers an agreed course of care to all expectant mothers living in the Republic of Ireland (HSE, 2021). However, there is no national database recording this care, as antenatal care is provided by the individual women's GPs and hospital obstetricians; therefore, recruiting those involved in this scheme would warrant extensive collaboration with antenatal service providers across Ireland, and would require access to fragmented healthcare records across different maternity units and GPs. As discussed in the previous section, one potential option could be collaborating with the MN-CMS Project, an ongoing endeavour to design and implement an EHR for all women and infants enrolled in maternity, newborn, and gynaecology services across Ireland. For sampling purposes, a mutually beneficial collaboration could possibly be established between the MN-CMS Project and the GUI study, whereby the two projects cooperate to get this system launched in the near future.

**Table 6.11** Options for the Irish context: prenatal sampling

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Prenatal sampling | Sampling during the prenatal period to study individuals from before birth. | • Gathers important prenatal information on the child's and mother's health during pregnancy.<br>• More accurate capture of critical events and stressors/resources during pregnancy.<br>• Information on biological mother for sample of children who are adopted at birth. | • Lack of centralised listings of pregnant women in Ireland (although this could change in the future). |

## 6.6.2 Ensuring representativeness

*Summary:* Samples must adequately represent members from traditionally hard-to-reach groups.

The robustness of the evidence produced by longitudinal birth cohorts is contingent on the ongoing representativeness of the target population (Leung *et al.*, 2013; Stafford *et al.*, 2013). Differential study participation, such as lower levels of participation among socially disadvantaged or ethnic minority groups, can prompt an underestimation of the actual prevalence of significant outcomes among these higher-risk groups, and limit the policy relevance of the study's findings (Woolfenden *et al.*, 2016). In order to adequately represent hard-to-reach groups, leading international birth cohort studies usually strive to recruit a sample size of at least 15,000 while oversampling relevant subpopulations. By combining a moderately large sample size with a targeted sampling strategy, these studies are more apt to generate valid findings on policy-relevant groups, study rare outcomes, and capitalise on novel methodologies and designs that necessitate larger sample sizes (Western *et al.*, 2014).

The first two decades of the new millennium have witnessed Ireland grow increasingly diverse as a country. Therefore, oversampling subpopulations of policy interest will be essential to the establishment of a credible longitudinal data resource for ethnic minority and disadvantaged families, alongside more advantaged families living in contemporary Ireland. According to the 2016 Census, there were 535,475 non-Irish nationals residing in Ireland, originating from 200 different nations (CSO, 2016b). In addition to enhancing Ireland's cultural diversity, immigrants to Ireland are highly diverse in terms of their age profile, socioeconomic status, culture, religion, language, and additional needs (e.g. disability, educational needs); this results in enormously varied experiences of Irish society across different groups (e.g. discrimination, employment opportunities, and access to public services such as healthcare and education) (McGinnity *et al.*, 2018). Alongside Ireland's rising immigrant population, the Traveller community is an extremely marginalised and traditionally underrepresented indigenous group. Travellers face extreme social exclusion and disadvantage, which is evident across the areas of education, healthcare, employment, and housing (Boyle, Flynn & Hanafin, 2020; Keogh *et al.*, 2020). Accordingly, adequate representativeness of the Traveller community and different immigrant populations via oversampling in a future Irish birth cohort study would be invaluable for public policy and service improvement.

Above all, the successful execution of this oversampling strategy would necessitate an in-depth knowledge of the relevant population subgroups and their geographical distribution, in addition to a sampling frame which contains sufficient detail about these indicators (Western *et al.*, 2014). Although sample stratification at the individual child/family level using socioeconomic and demographic variables may not be permissible under the GDPR, relevant data at an aggregate level (e.g. from the population census) could facilitate the identification of areas of higher ethnic minority concentration or social disadvantage across Ireland. For instance, since Traveller culture places a strong emphasis on living in close proximity to extended family, the population density of Travellers tends to be higher in certain cities (e.g. Cork, Galway) and towns (e.g. Tuam, Dundalk, Longford) throughout Ireland (CSO, 2016a). Therefore, oversampling children from the Traveller community would require stratification by geographical area and a more targeted strategy.

Besides using group statistics and known population figures to ensure representativeness, an analysis of achieved response rates during the pilot phases could also emphasise the need to oversample specific groups to guarantee a sufficient absolute number of participants for analysis. This was the case for the GUI Cohort '08, wherein observed pilot response rates for Cohort '08 at Wave 1 were significantly lower among unmarried mothers and non-national families. To address these higher levels of non-participation, unmarried mothers were slightly oversampled in the completed sample, whereas married mothers were undersampled by 6.1 percentage points (Thornton *et al.*, 2013). Similarly, to ensure adequate representation of non-national infants, a supplementary sample of 700 non-national families was included in accordance with external population estimates (Quail *et al.*, 2011a). The future birth cohort study could undergo piloting, simulate supplemental sample sizes needed, and conduct subsequent non-response bias analyses to ascertain the need to oversample relevant population groups and to determine the required sample size from each group.

**Table 6.12** Options for the Irish context: ensuring representativeness

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Piloting | A pilot study to analyse variation in response rates by geographical area and socioeconomic and demographic household characteristics. | • Detailed information regarding variation in response rates from the pilot phase helps determine the required sample size for each subgroup.<br>• At the pilot stage, response rates for different subpopulations could be checked against known population statistics to ensure representativeness. | • Since pilot studies tend to be based on small samples, they are subject to the limitations of small sample size analysis (e.g. limited statistical power; potentially higher variability). |

## 6.7 Data collection domains and measurements

*Summary:* A broad set of domains is of great interest to stakeholders and will enable the continuation of GUI study measures and topics to enable cross-cohort comparison. The advent of Europe's first birth cohort study, Growing Up In Digital Europe (GUIDE), offers a further avenue for data harmonisation in a new GUI birth cohort study.

In Chapter 3, we briefly examined the domains of data collection in the GUI study and in the selected named cohort studies. The GUI study is an interdisciplinary study that examines individual outcomes across the dimensions of physical activity and health, education and cognition, and socio-emotional and behavioural functioning, with contextual data collected on key sociodemographic indicators, lifestyle, families, housing quality, neighbourhoods, and schools. With a similar broad focus, the Millennium Cohort Study (MCS) captured data on individuals' physical, cognitive, socio-emotional, and behavioural functioning, and on contextual factors including family dynamics, housing, and neighbourhoods. In contrast, the Danish National Birth Cohort (DNBC) is more focused on data on health and disease, and precipitating factors. Étude Longitudinale Française depuis l'Enfance (ELFE), like the DNBC, collects data on health but as a demographic survey; it is also heavily focused on social dynamics and the environment. The MCS, DNBC, and ELFE have all collected biological data from their samples.

Stakeholders interviewed for this review expressed interest in a range of domains of data for a new GUI birth cohort study. Corresponding to domains covered in previous GUI waves, stakeholders wanted to see data collected on digital experiences and the online environment, on children's outdoor play and access to green areas, on children's housing quality and family structures, and on children's peer networks and friendships, and suggested using clustered sampling to better study the impact of neighbourhoods. Stakeholders also noticed a shift in early years education which they recommended to study in detail, alongside the suggestion to explore parental financial and employment circumstances. Stakeholders also proposed an examination of the role of Delivering Equality of Opportunity in Schools (DEIS) schools, alongside an investigation of the services available to children who are in poverty but who do not attend a DEIS school. Culture, race, and disability were also mentioned as important dimensions for a new GUI birth cohort study to cover. To obtain adequate coverage of minority groups, stakeholders recommended active recruitment via existing organisations and networks (e.g. Tusla, grassroots movements) which have established links with potential minority group participants.

When selecting specific measures to cover survey content domains, special attention must be given to measurement harmonisation in order to maximise the value of new data for impacting policy, practice, and research. Considerations for the GUI study include cross-cohort harmonisation, such that the same measurements can be used to generate comparable data across cohorts for research into time trends and history effects. There is also the issue of cross-study harmonisation – for example, with the new Early Life Cohort Feasibility Study in the UK. Harmonised measures can be selected based on existing study documentation, and in consultation with other cohort study designers.

GUIDE will be Europe's first cross-nationally comparative birth cohort study and is anticipated to launch in 2027, with pilot studies taking place across Europe in 2026. GUIDE plans to be an infant cohort study, with parent surveys conducted every 2 or 3 years for the first 8 years of the child's life. At age 8 years, children in GUIDE will be surveyed. There is also the potential for another cohort of 8-year-olds to be surveyed. Including this cohort could serve to complement the planned infant cohort, facilitating cross-age comparisons within the same historic period, and cross-era comparisons of children of the same age (Lynn, 2019). Currently, development work is under way to examine innovative modes of data collection, the potential for biological data, and data linkage. Like the GUI study and the MCS, GUIDE measurement domains cover children's physical, cognitive, emotional, behavioural, and social functioning. In addition, GUIDE captures parent and child well-being, taking the perspective that well-being is a synergistic indicator of all lived experiences (Szymczyk *et al.*, 2019). GUIDE also plans to measure the developing child's sociocultural context, including the role of culture, race, gender, and disability; the socialising impact of parents, teachers, and peers; and environmental factors, including housing, parental employment, neighbourhoods, digital lives, and climate change (Szymczyk *et al.*, 2019). Contingent on the future development and piloting of GUIDE survey instruments, a further consideration for a potential new GUI birth cohort is whether it might be possible to harmonise data measurement, content area, and data collection time points with those already planned for GUIDE in order to enable robust cross-cultural comparisons with other countries in Europe.

**Table 6.13** Options for the Irish context: data collection domains and measurements

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Cross-cohort harmonisation | Select domains and measures consistent with prior GUI cohorts. | • Cross-cohort comparability.<br>• Enables time trend analyses. | • The historic focus of GUI might not reflect current political priorities, social and economic change, or changed policy and service contexts. |
| Cross-study harmonisation | Select domains and measures consistent with other birth cohort studies internationally. | • Enables cross-national comparison. | • Different timing of data collection and sampling strategies across birth cohort studies reduces comparability. |
| Integration of GUIDE into the GUI study | Integrate the GUIDE cross-national study of child well-being with the GUI study. | • Could maximise data for reliable cross-national comparisons.<br>• Allows the evaluation of Irish outcomes against those in other countries to inform policy.<br>• Many of the GUIDE measures are taken from the GUI study, enabling some cross-cohort comparability.<br>• Would provide the opportunity to use all GUIDE measurements or to embed a core GUIDE module. | • Annual subscription fee would need to be paid for GUIDE hub so that Irish data can be incorporated into European-level reports.<br>• GUIDE alone is insufficient to generate specific policy-relevant data relevant to the Irish national context.<br>• The proposed timing of GUIDE may not match the developmental considerations which guide the data collection time frame of the GUI study. |

## 6.8 Data analysis and outputs

This section will examine data accessibility and ways to maximise data engagement via training, workshops, conferences, and international collaboration. Methods to track study impacts are also discussed.

### 6.8.1 Data access

*Summary:* Improvements in data management processes, security, and infrastructure make data more secure, open, and accessible.

Access to birth cohort datasets varies across studies, as demonstrated in Chapter 4 (see Section 4.5 Data access). For example, the DNBC datasets are not accessible in full; instead, selected variables are used as per the bespoke requirements of each secondary research study. Not only does this result in longer waiting times to access data, but it also results in dataset access and data handling fees (see Chapter 4, Appendix 4A for details on waiting times and fees to access birth cohort data).

It is possible to improve data management processes, security, and infrastructure to allow for more open and accessible data collaboration. There are good examples of networks that have created accessible databases and data infrastructures to increase data use and engagement. The EU Child Cohort Network, which is funded by the LifeCycle Project, is a resource that works around the findable, accessible, interoperable, and reusable (FAIR) principles to make birth cohort data findable, accessible, interoperable, and reusable (Jaddoe *et al.*, 2020). It integrates 19 pregnancy and child cohorts[54] (all of which are European, with the exception of one Australian cohort) and has harmonised and standardised data to allow for greater research analysis and collaboration at a large scale and at an individual level. The cohort data are oriented towards pre-conceptual, prenatal, postnatal, and early childhood exposures and their potential impact on physical (i.e. cardio, metabolic, and respiratory) and mental health outcomes. To ensure data security, harmonised

---

[54] Avon Longitudinal Study of Parents and Children (ALSPAC) (UK), ALSPAC-Generation 2 (UK), Born in Bradford (UK), EU Childhood Obesity Programme (Germany), DNBC (Denmark), the EDEN study (France), ELFE (France), Groningen Expert Center for Kids with Obesity (the Netherlands), Generation R (the Netherlands), Generation R *Next* (the Netherlands), Helsinki Birth Cohort Study (Finland), INMA Project - Children and the Environment (Spain), the Norwegian Mother and Child Cohort (MoBa) study (Norway), Nothern Finland Birth Cohort 1966 (Finland), Northern Finland Birth Cohort 1986 (Finland), the Nascita e INFanzia: gli Effetti dell’Ambiente (NINFEA) study (Italy), the Raine Study (Australia), the Rhea Study (Greece), and Southampton Women’s Survey (UK).

data are stored by the dedicated research institution in each cohort country, and can only be accessed by researchers through a shared, secure DataSHIELD platform.[55] This means that datasets do not have to be created for individual researchers or physically transferred, which results in more accessible data. The platform also works by recognising each cohort country's data regulations to allow for international collaboration (Jaddoe *et al.*, 2020).

Currently, COhort cOmmunity Research and Development Infrastructure Network for Access Throughout Europe (COORDINATE), funded by the European Commission under Horizon 2020, is developing a virtual portal for accessing studies that contain data on child well-being and that are archived in the Consortium of European Social Science Data Archives (CESSDA) (COORDINATE Network, 2022a). This will provide a portal of potentially hundreds of studies, including major cohort studies, which can help researchers identify and access data on children's development and experiences (COORDINATE Network, 2022a). For example, the GUI cohort study is held in the Irish Social Science Data Archive (ISSDA), which is linked to CESSDA. CESSDA's portal, developed as part of the COORDINATE Network, will make the GUI data more visible to scholars across Europe (COORDINATE Network, 2022b).

Another example is the UK Biobank, which has created an in-depth, accessible biomedical database for national and international researchers. To utilise the data, researchers must register with the UK Biobank, meet registration protocols, and wait approximately 10 working days for approval. A fee is in place for servicing a data application and there is a tiered fee depending on the level of data access required.[56]

Maximising the use of the wealth of data collected by current and future GUI studies is crucial. The GUI study has a well-regarded history of making Anonymised Microdata File (AMF) datasets widely available and accessible to researchers in Ireland, and this should continue. However, GUI data could be more widely accessible to improve cross-country comparison (which would benefit policy evaluation (see Mazzonna, 2014)), and to better engage with globalised research collaboration

---

[55] DataSHIELD is an infrastructure and series of R packages that enables the remote and non-disclosive analysis of sensitive research data; see https://www.datashield.org/.

[56] The UK Biobank registration protocols can be found at https://www.ukbiobank.ac.uk/enable-your-research/register.

(Bardosh *et al.*, 2020). International use of more sensitive GUI data is currently unfeasible for many researchers, as the GUI study's terms and conditions state that requests for the provision of Researcher Microdata File (RMF) datasets to or from locations outside of the Republic of Ireland will generally not be facilitated unless researchers are based in Ireland for data analysis. This makes it more challenging for international researchers to engage with GUI data or to engage in cross-country analysis of sensitive variables such as ethnicity and psychopathology. To facilitate cross-country comparison, a harmonised 'Growing Up' database could potentially be created to facilitate international comparisons between the various 'Growing Up' studies, including Growing Up in Ireland, Growing Up in Scotland, Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC), and Growing Up in New Zealand.

**Table 6.14** Options for the Irish context: data analysis

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Harmonised 'Growing Up' international database | A database including harmonised data from all 'Growing Up' studies (i.e. Ireland, Scotland, Australia, and New Zealand). | • Cross-country comparative analysis using similar (i.e. harmonised) measures would produce valid and policy-relevant findings. | • Post-hoc harmonisation is labour intensive and subject to limitations of the available data. Some measures are not collected in every study; others are collected using different instruments. |
| Make RMFs accessible to non-Ireland-based researchers. | GUI RMF datasets include more detailed and sensitive data than the AMF datasets and are currently not accessible to researchers outside of Ireland. | • Removing the location restriction would likely increase GUI data use and result in more cross-country comparative analyses. | • Given the current approval process that requires intensive involvement of the CSO, removing location restrictions would entail staffing and cost implications for the CSO. |

## 6.8.2 Maximising data engagement

As addressed in Section 4.5 of Chapter 4, there is a mixture of readily available and harder to find information regarding the types of training, workshops or conferences associated with data from the named birth cohort studies (i.e. DNBC, ELFE). To maximise the use of birth cohort data available for secondary analysis, knowledge sharing and accessible supports must be prioritised. In order to target information to the appropriate channels and users, and to gauge its impact, it is recommended that a future GUI birth cohort study develops evidence-based strategies for maximising data engagement, and regularly evaluates whether these provided resources and training have a positive influence on user engagement.

A barrier to engagement with the GUI study was noted in the stakeholder consultation by stakeholders working in non-profit and government-affiliated sectors. Those stakeholders spoke of needing additional resources to use GUI data because they did not have the data analysis expertise or the staffing necessary to produce their own reports using GUI data. To encourage the use of GUI data, further partnerships between the GUI study team and a wide range of industry stakeholders is recommended. A further option could be to partner with and support higher education programmes that focus on complex and novel longitudinal research methodologies and data analysis. In addition, stakeholders from non-profit organisations highlighted the need to provide targeted training in GUI data analysis.

Collaborations between research teams and communities tend to produce the highest-impact scientific work (Leone Sciabolazza *et al.*, 2017). The stakeholders suggested that to maximise engagement with the GUI study, and to make the study even more robust in terms of its design and content, community groups and citizens should be more centrally involved in the design and analysis of GUI data. One consideration is to create a GUI programme for citizen science, wherein communities and individuals are assisted to analyse the GUI data – perhaps in order to solve key social issues. This type of model could build on initiatives for engaging citizens in science, such as the Government of Ireland's 'Creating Our Future' campaign, which collected thousands of ideas from citizens on research priorities for Ireland across 2021 and 2022. Funding could be made available for community groups, supported by researchers, to use GUI data to analyse topics of significant interest that emerged across the Government of Ireland consultation. A further suggestion is to routinely offer opportunities for community organisations and HEIs to collaborate together to analyse GUI data for the benefit of the community organisations. This would address the issue raised by stakeholders regarding a lack of internal scientific expertise to make the best use of GUI data.

**Table 6.15** Options for the Irish context: maximising engagement

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Frequent and regular data user workshops and conferences | A new GUI study would build on the user engagement strategy of the current GUI study (e.g. annual conference) and offer data user workshops at more frequent intervals. | • Ongoing engagement with potential and existing data users helps promote the wide-ranging use of data for academic and policy-relevant research. | • There are additional costs associated with frequent and regular workshops. |
| GUI citizen science programme | A programme to encourage citizens and community groups to become involved in the design and analysis of the GUI study and resulting data. | • Engages citizens and community groups.<br>• Enables more use of the data.<br>• Enhances the national reputation of the GUI study.<br>• Could act as an example of best practice for international showcasing.<br>• Provides the opportunity to involve national research funders for citizen science. | • Financial and administrative costs involved. |

## 6.8.3 Training

To support the use of GUI data, GUI researchers host regular data workshops, with the materials being made available on the GUI website afterwards.[57] These workshops are designed for data users with no previous experience with GUI data; thus, they involve a general preparatory focus. In addition, GUI researchers host a free-to-attend annual research conference that offers researchers and academics the opportunity to present their papers based on GUI data. Both the workshops and the

---

[57] For information about GUI data workshops, see https://www.growingup.ie/information-for-researchers/growing-up-in-ireland-data-workshop/.

annual conference are valuable data training and data engagement resources which should remain in place for a future GUI birth cohort.

Given the complex nature of longitudinal data, in order to maximise the data collected in a future GUI study, further training could be developed for GUI data users to incorporate advanced data analysis methods, including already established as well as novel or dynamic methods (Wooldridge, 2016; Angrist & Pischke, 2009). At the centre of any new training should be the awareness that researchers at every stage of their careers, from PhD students to senior researchers, need to update and inform their skills regarding new methodological approaches and ongoing methodological developments. Depending on the data collection design of a future GUI birth cohort, this could include not only advanced statistical techniques for panel data analysis, but also training on data linkage (e.g. how to deal with linkage error) (Harron, Doidge & Goldstein, 2020; Harron *et al.*, 2017; Gustavson *et al.*, 2012). To build upon training and to promote engagement with the data, advice hubs could be created for a future GUI birth cohort as a consultation resource for researchers and stakeholders who use GUI data.

The Centre for Longitudinal Studies in Ireland (CLSI)[58] and CLOSER[59] both work towards increasing awareness of the effectiveness of longitudinal research data. These organisations also promote best practices in working with longitudinal research data, and with both qualitative and quantitative longitudinal data. CLOSER frequently runs training workshops for researchers and data users at all stages of their careers, and for those working in practice and policy settings. It also provides free online material via the CLOSER Learning Hub, using real-life anonymised cases, to help students and data analysts better understand longitudinal analysis. Similarly, the CLSI has run a Spotlight Series for data users to learn more about longitudinal research. However, this more advanced (academic focused) training for longitudinal studies is not currently available through other venues in Ireland. An ideal future scenario would be to partner with HEIs to help design and administer qualifications

---

[58] For more information about the CLSI, visit https://www.clsi.ie/.

[59] For more information about CLOSER UK, visit https://www.closer.ac.uk/.

in secondary data analysis and cohort study methods to support the next generation of GUI cohort study scientists and data analysts.

Additionally, the importance of networks to maximise data usage should be exploited (Kyvik & Reymert, 2017). It can be a challenge for study teams to reach their target groups, such as PhD students or community stakeholders. An information promotion strategy involving outreach to various stakeholders across looser networks would be valuable for a future cohort study. Stakeholder participants also identified this need and recommended building closer relationships to increase the knowledge and value of the GUI study in specific cultural communities.

**Table 6.16** Options for the Irish context: training

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Training workshops and free online material | A series of data and methods training workshops, with materials posted for free online. | • Accessible workshops and training materials would benefit a diverse audience of early career and experienced researchers working in both academic and non-academic settings. | • There are additional costs associated with methods training workshops. |
| Higher education qualification in cohort studies | Support for a master's programme or other form of qualification in quantitative social sciences/cohort studies with a GUI focus. | • Develops the next generation of cohort study scientists in Ireland – potential GUI staff and data users.<br>• Would have value for international students.<br>• Pays for itself once student intake is secured. | • There are administrative costs of developing a new programme in partnership with a higher education institution. |

### 6.8.4 International collaborations

It is also recommended that a future GUI study develop a strategy to promote GUI research data internationally through collaborations with other cohort networks (e.g.

CLOSER, the Society for Longitudinal and Lifecourse Studies),[60] and international birth cohort studies such as the 'Growing Up' studies in Australia, New Zealand and Scotland. Moreover, international data engagement could be streamlined if a secure, shared, harmonised dataset of interlinking features were available across different cohort studies (Wey *et al.*, 2021), including the GUI studies. Furthermore, a future GUI birth cohort study could help to facilitate transnational access visits for international researchers to analyse GUI data in Ireland – for example, by funding hosting visits awarded to research institutions or non-profit organisations. This strategy is currently being used by the COORDINATE Network[61] to facilitate access to the GUI data in Ireland, as well as to 10 other large-scale cohort studies across Europe. Access visits are awarded on a competitive basis and facilitate data access and analysis support from professionals who regularly use GUI and other cohort study datasets.

### 6.8.5 Track the GUI study's impact

Based on the work undertaken for this review, it was noted that secondary papers associated with the named birth cohort studies can be difficult to track. For example, this research team identified that secondary research papers often do not outline in their abstracts the cohort study name or dataset(s) used. Researchers who publish studies using GUI data should ensure that reports and publications are easily searchable through the use of appropriate keywords, well-considered abstracts, and the use of unique DOIs for each dataset which can be cited in other literature.

**Table 6.17** Options for the Irish context: dataset DOIs

| Option | Description | Advantages | Disadvantages |
|---|---|---|---|
| Unique DOI for each dataset | Each GUI dataset would have its own unique DOI. | • DOIs make it easier to discover and cite surveys.<br>• DOIs ensure consistency and accuracy of citations.<br>• DOIs enhance stability, providing a permanent and persistent link to GUI datasets even if GUI-related websites disappear. | • There are different DOI providers, some free and others charging fees. |

---

[60] For more information about the Society for Longitudinal and Lifecourse Studies in the UK, visit https://www.slls.org.uk/.

[61] For more information about the COORDINATE Network, visit https://www.coordinate-network.eu/about.

## 6.9 Conclusion: rationale for a new GUI birth cohort study

Looking ahead into the remaining 2020s, we can see both a tunnel and a wide horizon. Irish society is structured in such a way that social systems maintain inequalities and foster predictable lived experiences and developmental outcomes. Yet, changes that no one can predict (such as the COVID-19 pandemic or the war in Ukraine) can rupture or consolidate these structures, more clearly revealing the structural causes of human growth and offering new pathways towards human flourishing. A new GUI birth cohort can observe these continuities and discontinuities in society and individual development, tracking the impact of policies and other initiatives across the life course. Only a longitudinal study can achieve this level of clarity in explaining cause and effect and is best placed to respond to structural changes with an involved sample of individuals who have prior data histories ready to examine.

Ireland's social demography is continuously changing, with numbers of immigrants being relatively even to numbers of emigrants (CSO, 2021b), and with most migrants being non-Irish nationals (CSO, 2018). Some schools now have majority ethnic minority and migrant populations, although teachers are predominantly White Irish (Devine *et al.*, 2020). There is also the opportunity for more diverse household compositions following the legalisation of same-sex marriage in Ireland in 2015, and with immigrant families bringing with them cultural differences in how families live together. Furthermore, there is an increased policy focus on the inclusion of traditionally marginalised groups, such as persons living with disability[62] and the Traveller/Roma community.[63] A new GUI birth cohort is needed to sample this new and diverse Irish population, and to trace the lives of these individuals in social, economic, and political contexts.

A new GUI birth cohort can also take advantage of methodological advances in data collection that include online questionnaires, wearables, and smartphone apps. Capturing real-time data on mothers' physical activity and stress levels – both during and after pregnancy – could advance our understanding of child development not

---

[62] *National Disability Inclusion Strategy 2017-2021*, https://www.justice.ie/en/JELR/dept-justice-ndi-inclusion-stratgey-booklet.pdf/Files/dept-justice-ndi-inclusion-stratgey-booklet.pdf

[63] *National Traveller and Roma Inclusion Strategy 2017 – 2021*, https://www.justice.ie/en/JELR/National%20Traveller%20and%20Roma%20Inclusion%20Strategy,%202017-2021.pdf/Files/National%20Traveller%20and%20Roma%20Inclusion%20Strategy,%202017-2021.pdf

only in Ireland but also internationally. Also, parents can use smartphone apps to record information on their child's behaviour and cognitive and emotional functioning. Apps can prompt parents to record their child's sleep patterns and to test their child without a researcher present in order to provide in situ, naturalistic measurements of children's cognitive and behavioural functioning. These advances can provide valuable information collected at shorter time intervals and at relatively low cost, supplementing fieldwork interviews administered once every few months or years. A new GUI birth cohort study could add these methodologies to its toolkit, providing for an enriched, world-leading understanding of child development and well-being.

The COVID-19 pandemic is still with us and may continue long into the 2020s. Changes in the population's everyday health are impacting the economy and societal functioning, and the consequences of 'long COVID' are still in their infancy. A new GUI birth cohort would be ideal to study the impact of the COVID-19 pandemic on infants, and on parents of infants, who may be affected by economic and health issues relating to the pandemic. Although it is possible to study the current and future impact of the pandemic with cross-sectional surveys, only a longitudinal study like the GUI study can rigorously unpack the longer-term influence of key pandemic risk factors, including changes in financial status, increased cost of living, parental stress, parental employment and work–life balance, and school experiences (e.g. senior cycle reform), by accurately measuring these factors at the time of exposure.

Finally, a new GUI birth cohort study will allow for consideration of a plan for data harmonisation with existing GUI cohorts, and with other 'Growing Up' studies, such as the new European birth cohort study, GUIDE. Data harmonisation efforts have developed as the GUI study has matured, progressing from the first few waves of the survey, when this longer-term planning for harmonisation was underutilised. A new GUI birth cohort study can yield more valuable comparisons through careful data harmonisation, allowing us to access information about changes in Irish society and about how Irish children and parents are faring in comparison with those across Europe. This increased informational power should be a worthwhile investment and highlights the value of data linkage nationally (via the CSO) and across the rest of Europe.

# References

Aarts, N., Bunnik, E. M. and Boeckhout, M. (2017) *Guide to the detection, management and communication of incidental findings for biobanks in BBMRI-NL*. Rotterdam: Erasmus University Medical Center.

Ajslev, T. A., Andersen, C. S., Gamborg, M., Sørensen, T. I. and Jess, T. (2011) 'Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics', *International Journal of Obesity*, 35(4), pp. 522–529. Available at: https://doi.org/10.1038/ijo.2011.27 (Accessed: 28 February 2022).

Al Baghal, T. (2016) 'Obtaining data linkage consent for children: Factors influencing outcomes and potential biases', *International Journal of Social Research Methodology*, 19(6), pp. 623–643. Available at: https://doi.org/10.1080/13645579.2015.1064635 (Accessed: 28 February 2022).

Al-Rifai, R. H., Nasloon, A., Barigye, E. T., Al Haddad, A. H. I., Al-Maskari, F., Loney, T. and Ahmed, L. A. (2020) 'Maternal and birth cohort studies in the Gulf Cooperation Council countries: A systematic review and meta-analysis', *Systematic Reviews*, 9, pp. 1–13. Available at: https://doi.org/10.1186/s13643-020-1277-0 (Accessed: 28 February 2022).

Allen, M. (2017) 'Sampling Frames', in *The SAGE Encyclopedia of Communication Research Methods*. Milwaukee: SAGE Publications, Inc. Available at: https://doi.org/10.4135/9781483381411.n543 (Accessed: 28 February 2022).

Almond, D. and Currie, J. (2011) 'Killing Me Softly: The Fetal Origins Hypothesis', *The Journal of Economic Perspectives,* 25(3), pp. 153–172. Available at: https://doi.org/10.1257/jep.25.3.153 (Accessed: 28 February 2022).

Altman, R., Lee, B., Lichtveld, M., Miranda, M. L., Perry, C. and Zoghbi, H. (2014) *National Children's Study (NCS) Working Group Final Report*. US: National Institutes of Health.

Andersen, A., and Kantsø, L. S. (2018a) *Code book for 3rd interview (the child is 6 months old.* Section of Epidemiology, Institute of Public Health, University of Copenhagen. https://www.dnbc.dk/-/media/arkiv/projekt-sites/dnbc/kodeboeger/interviews-1-4/code_book_interview_3.pdf

Andersen, A., and Kantsø, L. S. (2018b) *Code book for 4th interview (the child is 18 months old).* Copenhagen: Section of Epidemiology, Institute of Public Health, University of Copenhagen. https://www.dnbc.dk/-/media/arkiv/projekt-sites/dnbc/kodeboeger/interviews-1-4/code_book_interview_4.pdf

Andersen, A., Larsen, P. S. and Kantsø, L. S. (2018a) *Code Book for 1st Interview (12th gestation week).* Copenhagen: Section of Epidemiology, Institute of Public Health, University of Copenhagen. Available at: https://www.dnbc.dk/-/media/arkiv/projekt-sites/dnbc/kodeboeger/interviews-1-4/code-book-interview-1.pdf?la=en (Accessed: 28 February 2022).

Andersen, A., Larsen, P. S. and Kantsø, L. S. (2018b) *Code book for the 2nd interview (30th week of gestation)*. Copenhagen: Section of Epidemiology, Institute of Public Health, University of Copenhagen. Available at: https://www.dnbc.dk/-/media/arkiv/projekt-sites/dnbc/kodeboeger/interviews-1-4/code_book_interview_2.pdf?la=en (Accessed: 28 February 2022).

Anderson, S., Bradshaw, P., Cunningham-Burley, S., Hayes, F., Jamieson, L., MacGregor, A., Marryat, L. and Wasoff, F. (2007) *Growing Up in Scotland: A Study Following the Lives of Scotland's Children*. Edinburgh: Scottish Executive. Available at: http://www.scotland.gov.uk/Publications/2007/01/17162004/0 (Accessed: 28 February 2022).

Angrist, J. D., and Pischke, J.-S. (2009). *Mostly harmless econometrics.* Princeton: Princeton University Press.

Araujo, W. R. M., Santos, I. S., Barros, F. C. and Matijasevich, A. (2018) 'Birth Cohort Studies and the Early Origins of Adult Health and Disease', *NeoReviews*, 19(6), pp. e313–e321. Available at: https://doi.org/10.1542/neo.19-6-e313 (Accessed: 28 February 2022).

Ashing, K. T., Fernandez, A., Serrano, M., Garcia, M. and Nerio, K. (2018) 'Abstract 5287: Challenges and opportunities using mobile technology for data collection in biomedical research: An observational discovery science investigation with Spanish-speaking Latinas', *Cancer Research*, 78(13 Supplement), pp. 5287–5287. Available at: https://doi.org/10.1158/1538-7445.AM2018-5287 (Accessed: 28 February 2022).

Australian Institute of Health and Welfare (2021) *Australia's Children.* Cat. No. CWS 69. Canberra: Australian Institute of Health and Welfare. Available at: https://www.aihw.gov.au/ (Accessed: 28 February 2022).

Bailey, S. R., Townsend, C. L., Dent, H., Mallet, C., Tsaliki, E., Riley, E. M., Noursadeghi, M., Lawley, T. D., Rodger, A. J., Brocklehurst, P. and Field, N. (2017) 'A pilot study to understand feasibility and acceptability of stool and cord blood sample collection for a large-scale longitudinal birth cohort', *BMC Pregnancy and Childbirth*, 17. Available at: https://doi.org/10.1186/s12884-017-1627-7 (Accessed: 28 February 2022).

Bales, M., Pambrun, E., Melchior, M., Glangeaud-Freudenthal, N. C., Charles, M. A., Verdoux, H. and Sutter-Dallay, A. L. (2015) 'Prenatal psychological distress and access to mental health care in the ELFE cohort', *European Psychiatry*, 30(2), pp. 322–328. Available at: https://doi.org/10.1016/j.eurpsy.2014.11.004 (Accessed: 28 February 2022).

Bandara, D., Edwards, B., Mohal, J. and Daraganova, G. (2019) *Consent to data linkage in a child cohort study, Growing Up in Australia: The Longitudinal Study of Australian Children.* CSRM Working Paper No. 7/2019. Canberra: ANU Centre for Social Research & Methods Research, School of Social Sciences, The Australian National University.

Bandyopadhyay, A., Tingay, K., Akbari, A., Griffiths, L., Cortina-Borja, M., Bedford, H., Walton, S., Dezateux, C. and Lyons, R. A. (2018) 'Integrating electronic health records from different sources across the UK: lessons from a record linkage study', *International Journal of Population Data Science*, 3(2). Available at: https://doi.org/10.23889/ijpds.v3i2.495 (Accessed: 28 February 2022).

Bardosh, K. L., de Vries, D. H., Abramowitz, S., Thorlie, A., Cremers, L., Kinsman, J. and Stellmach, D. (2020) 'Integrating the social sciences in epidemic preparedness and response: A strategic framework to strengthen capacities and improve Global Health security', *Globalization and Health*, 16(1), p. 120. Available at: https://doi.org/10.1186/s12992-020-00652-6 (Accessed: 28 February 2022).

Barker, D. J. (2001) 'Fetal and infant origins of adult disease', *Monatsschrift Kinderheilkunde*, 149(1), pp. S2–S6. Available at: https://doi.org/10.1007/s001120170002 (Accessed: 28 February 2022).

Barker, D. J. P. (1995) 'Fetal origins of coronary heart disease', *BMJ*, 311(6998), pp. 171–174. Available at: https://doi.org/10.1136/bmj.311.6998.171 (Accessed: 28 February 2022).

Barker, D. J. P. (2007) 'The origins of the developmental origins theory', *Journal of Internal Medicine*, 261(5), pp. 412–417. Available at: https://doi.org/10.1111/j.1365-2796.2007.01809.x (Accessed: 28 February 2022).

Barker, D. J. P. and Osmond, C. (1986) 'Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales', *The Lancet*, 327(8489), pp. 1077–1081. Available at: https://doi.org/10.1016/S0140-6736(86)91340-1 (Accessed: 28 February 2022).

BASELINE (n.d.) *The Cork BASELINE study: Patient Information*. Available at: http://www.baselinestudy.net/ (Accessed: 28 February 2022).

Ben-Shlomo, Y. and Kuh, D. (2002) 'A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives', *International Journal of Epidemiology*, 31(2), pp. 285–293. Available at: https://doi.org/10.1093/ije/31.2.285 (Accessed: 28 February 2022).

Béranger, R., Hardy, E. M., Dexet, C., Guldner, L., Zaros, C., Nougadère, A., Metten, M. A., Chevrier, C., and Appenzeller, B. M. R. (2018) 'Multiple pesticide analysis in hair samples of pregnant French women: results from the ELFE national birth cohort', *Environment International*, 120, pp. 43–53. Available at: https://doi.org/10.1016/j.envint.2018.07.023 (Accessed: 28 February 2022).

Berthomier, N. and Octobre, S. (2019) 'The ELFE study of screen use in children aged from 0 to 2 years', *Culture études*, 1, pp. 1–32. Available at: https://doi.org/10.3917/cule.191.0001 (Accessed: 28 February 2022).

Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J., Devercelli, A. E., Wodon, Q. T., Vargas-Barón, E. and Grantham-McGregor, S. (2017) 'Early childhood development coming of age: Science through the life course', *The Lancet*, 389(10064), pp. 77–90. Available at: https://doi.org/10.1016/S0140-6736(16)31389-7 (Accessed: 28 February 2022).

Blane, D., Kelly-Irving, M., d'Errico, A., Bartley, M. and Montgomery, S. (2013) 'Social-biological transitions: How does the social become biological?', *Longitudinal and Life Course Studies*, 4(2), pp. 136–146. Available at: https://doi.org/10.14301/llcs.v4i2.236 (Accessed: 28 February 2022).

Bliddal, M., Liew, Z., Pottegård, A., Kirkegaard, H., Olsen, J. and Nohr, E. A. (2018) 'Examining Nonparticipation in the Maternal Follow-up Within the Danish National Birth Cohort', *American Journal of Epidemiology*, 187(7), pp. 1511–1519. Available at: https://doi.org/10.1093/aje/kwy002 (Accessed: 28 February 2022).

Blumenberg, C., Zugna, D., Popovic, M., Pizzi, C., Barros, A. J. D. and Richiardi, L. (2018) 'Questionnaire Breakoff and Item Nonresponse in Web-Based Questionnaires: Multilevel Analysis of Person-Level and Item Design Factors in a Birth Cohort', *Journal of Medical Internet Research*, 20(12), 1-11.Available at: https://doi.org/10.2196/11046 (Accessed: 28 February 2022).

Bone, J. N., Pickerill, K., Kinshella, M.-L. W., Vidler, M., Craik, R., Poston, L., Stones, W., Sevene, E., Temmerman, M., Etyang, A. K., Roca, A., Russell, D., Tribe, R. M., von Dadelszen, P. and Magee, L. A. (2020) 'Pregnancy cohorts and biobanking in sub-Saharan Africa: A systematic review', *BMJ Global Health*, 5(11), p. e003716. Available at: https://doi.org/10.1136/bmjgh-2020-003716 (Accessed: 28 February 2022).

Boyle, A., Flynn, M. and Hanafin, J. (2020) 'Optimism despite disappointment: Irish Traveller parents' reports of their own school experiences and their views on education', *International Journal of Inclusive Education*, 24(13), pp. 1389–1409. Available at: https://doi.org/10.1080/13603116.2018.1530805 (Accessed: 28 February 2022).

Bradshaw, P. and Hinchliffe, S. (2018) *Growing Up in Scotland: Overweight and Obesity at Age 10*. Edinburgh: Scottish Centre for Social Research. Available at: https://www.gov.scot/binaries/content/documents/govscot/publications/research-and-analysis/2018/11/growing-up-scotland-overweight-obesity-age-10/documents/00543330-pdf/00543330-pdf/govscot%3Adocument/00543330.pdf (Accessed: 28 February 2022).

Bradshaw, P., Bromley, C., Hill., T., Mabelis, J., Parkes, A., Smith, K., Sweeting, H., Warner, P. and Wright., D. (2013) *Growing Up in Scotland: Birth cohort 2: results from the first year*. Edinburgh: Scottish Government. Available at: http://www.nls.uk/scotgov/2013/9781782563617.pdf (Accessed: 28 February 2022).

Brandstetter, S., Toncheva, A. A., Niggel, J., Wolff, C., Gran, S., Seelbach-Göbel, B., Apfelbacher, C., Melter, M. and Kabesch, M. (2019) 'KUNO-Kids birth cohort study: Rationale, design, and cohort description', *Molecular and Cellular Pediatrics,* 6, p. 1. Available at: https://doi.org/10.1186/s40348-018-0088-z (Accessed: 28 February 2022).

Bray, I., Noble, S., Robinson, R., Molloy, L. and Tilling, K. (2017) 'Mode of delivery affected questionnaire response rates in a birth cohort study', *Journal of Clinical Epidemiology*, 81, pp. 64–71. Available at: https://doi.org/10.1016/j.jclinepi.2016.09.004 (Accessed: 28 February 2022).

Briody, J. (2021) 'Parental unemployment during the Great Recession and childhood adiposity', *Social Science and Medicine*, 275, 1-10. Available at: https://doi.org/10.1016/j.socscimed.2021.113798 (Accessed: 28 February 2022).

Bromley, C. and Cunningham-Burley, S. (2010) *Growing Up in Scotland: Health inequalities in the early years*. Edinburgh: Scottish Government. Available at: http://www.nls.uk/scotgov/2010/GUSresearchfindings2.pdf (Accessed: 28 February 2022).

Bronfenbrenner, U. (1979) *The ecology of human development: Experiments by nature and design.* Cambridge, Massachusetts: Harvard University Press.

Bronfenbrenner, U. (1994) 'Ecological models of human development', *Readings on the Development of Children,* 2(1), pp. 37–43.

Bronfenbrenner, U. and Morris, P. A. (2006) 'The Bioecological Model of Human Development', in Lerner, R. M. and Damon, W. (eds.) *Handbook of child psychology: Theoretical models of human development*. Hoboken, New Jersey: John Wiley & Sons Inc, pp. 793–828.

Burg, T., Kowarik, A., Six, M., Brancato, G. and Krapavickaité, D. (2019) *Quality Guidelines for Frames in Social Statistics (ESSnet KOMUSO Quality in Multisource Statistics)*. EUROSTAT. Available at: https://ec.europa.eu/eurostat/cros/system/files/qgfss-v1.51.pdf (Accessed: 28 February 2022).

Burke, V., Beilin, L. J., Simmer, K., Oddy, W. H., Blake, K. V., Doherty, D., Kendall, G. E., Newnham, J. P., Landau, L. I. and Stanley, F. J. (2005) 'Breastfeeding and Overweight: Longitudinal Analysis in an Australian Birth Cohort', *The Journal of Pediatrics*, 147(1), pp. 56–61. Available at: https://doi.org/10.1016/j.jpeds.2005.03.038 (Accessed: 28 February 2022).

Cai, M., Tan, K. H. and Ang, S. B. (2019) 'I-ACT: Integrated study on effect of Activity on ComplicaTions in pregnancy: study protocol of a multiethnic prospective cohort study', *BMJ Open*, 9(4), p. e025970. Available at: https://doi.org/10.1136/bmjopen-2018-025970 (Accessed: 28 February 2022).

Calderwood, L., Gilbert, E., Bann, D., Patalay, P. and Sullivan, A. (2019) *Survey methods research: New technologies*. London: Centre for Longitudinal Studies. Available at: https://esrc.ukri.org/files/funding/funding-opportunities/executive-summary-survey-methods-research-on-new-technologies/ (Accessed: 28 February 2022).

Camier, A., Chabanet, C., Davisse-Paturet, C., Ksiazek, E., Lioret, S., Charles, M.-A., Nicklaus, S. and de Lauzon-Guillain, B. (2020) 'Characterization of Infant Feeding Practices and Related-Family Characteristics in the French Nationwide ELFE Birth Cohort', *Nutrients*, 13(1), p. 33. Available at: https://doi.org/10.3390/nu13010033 (Accessed: 28 February 2022).

Canova, C. and Cantarutti, A. (2020) 'Population-Based Birth Cohort Studies in Epidemiology', *International Journal of Environmental Research and Public Health*, 17(15), p. 5276. Available at: https://doi.org/10.3390/ijerph17155276 (Accessed: 28 February 2022).

Casteel, A. and Bridier, N. (2021) 'Describing Populations and Samples in Doctoral Student Research', *International Journal of Doctoral Studies*, 16, pp. 339–362. Available at: https://doi.org/10.28945/4766 (Accessed: 28 February 2022).

Cawley, C., Buckenmeyer, H., Jellison, T., Rinaldi, J. B. and Vartanian, K. B. (2020) 'Effect of a Health System-Sponsored Mobile App on Perinatal Health Behaviors: Retrospective Cohort Study', *JMIR MHealth and UHealth*, 8(7), p. e17183. Available at: https://doi.org/10.2196/17183 (Accessed: 28 February 2022).

Cernat, A. and Sakshaug, J. W. (2021) 'Estimating the Measurement Effects of Mixed Modes in Longitudinal Studies: Current Practice and Issues', in P. Lynn (ed.) *Advances in Longitudinal Survey Methodology.* Essex: John Wiley & Sons, pp. 235–239.

Charles, M. A., Thierry, X., Lanoe, J.-L., Bois, C., Dufourg, M.-N., Popa, R., Cheminat, M., Zaros, C. and Geay, B. (2020) 'Cohort Profile: The French national cohort of children (ELFE): birth to 5 years', *International Journal of Epidemiology*, 49(2), pp. 368–369. Available at: https://doi.org/10.1093/ije/dyz227 (Accessed: 28 February 2022).

Charles, M., Leridon, H., Dargent, P. and Geay, B. (2011) 'Tracking the lives of 20,000 children. Launch of the Elfe child cohort study', *Population & Societies*, 475(2), pp. 1–4.

Checkoway, H., Pearce, N. and Kriebel, D. (2007) 'Selecting appropriate study designs to address specific research questions in occupational epidemiology', *Occupational and Environmental Medicine*, 64(9), pp. 633–638. Available at: https://doi.org/10.1136/oem.2006.029967 (Accessed: 28 February 2022).

Christensen, D., Fahey, M. T., Giallo, R. and Hancock, K. J. (2017) 'Longitudinal trajectories of mental health in Australian children aged 4-5 to 14-15 years', *PLOS ONE*, 12(11), p. e0187974. Available at: https://doi.org/10.1371/journal.pone.0187974 (Accessed: 28 February 2022).

Citizens Information (2021a) *Child Benefit*. Available at: https://www.citizensinformation.ie/en/social_welfare/social_welfare_payments/social_welfare_payments_to_families_and_children/child_benefit.html#l4c076 (Accessed: 28 February 2022).

Citizens Information (2021b) *Registering the birth of your baby*. Available at: https://www.citizensinformation.ie/en/birth_family_relationships/after_your_baby_is_born/registering_birth_your_baby.html (Accessed: 28 February 2022).

Clarke, N., Vale, G., Reeves, E. P., Kirwan, M., Smith, D., Farrell, M., Hurl, G. and McElvaney, N. G. (2019) 'GDPR: An impediment to research?', *Irish Journal of Medical Science,* 188(4), pp. 1129–1135. Available at: https://doi.org/10.1007/s11845-019-01980-2 (Accessed: 28 February 2022).

CLS (2019) *Millennium Cohort Study: A guide to the linked education administrative datasets.* 1st edn. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2019/07/User-guide-to-the-Linked-Education-Datasets-2019-Version2_May2019.pdf (Accessed: 28 February 2022).

CLS (2020) *Millennium Cohort Study: User Guide (Surveys 1 – 5). London:* Centre for Longitudinal Studies, Institute of Education, University of London. Available at: http://doc.ukdataservice.ac.uk/doc/5795/mrdoc/pdf/mcs1-5_user_guide_ed9_2020-08-07.pdf (Accessed: 28 February 2022).

CLS (2021a) *CLS Millennium Cohort Study: MCS Data Sweeps*. London: Centre for Longitudinal Studies. Available at: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/ (Accessed: 28 February 2022).

CLS (2021b) *Early Life Cohort Feasibility Study: Non-questionnaire measures, and additional informant interview topics*. London: Centre for Longitudinal Studies, University College London. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/02/ELCFS_Non-questionnaire-_ForConsultation.pdf (Accessed: 28 February 2022).

CLS (n.d.) *Millennium Cohort Study, MCS Sweeps*. Available at: https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/ (Accessed: 28 February 2022).

Cocchi, D., Giovinazzi, F. and Lynn, P. (2019) *Handling Non-Response in Longitudinal Surveys: A Review of Minimization and Adjustment Strategies and Implications for EuroCohort.* EuroCohort Working Paper No. 1. Daugavpils, Latvia: Daugavpils University. Available at: http://www.eurocohort.eu/ (Accessed: 28 February 2022).

Cognez, N., Warembourg, C., Zaros, C., Metten, M. A., Bouvier, G., Garlantézec, R., Charles, M. A., Béranger, R. and Chevrier, C. (2019) 'Residential sources of pesticide exposure during pregnancy and the risks of hypospadias and cryptorchidism: the French ELFE birth cohort', *Occupational and Environmental Medicine*, 76(9), pp. 672–679. Available at: https://doi.org/10.1136/oemed-2019-105801 (Accessed: 28 February 2022).

Collins, P. M. and Marassi, S. (2021) ‘Is That Lawful? Data Privacy and Fitness Trackers in the Workplace’, *International Journal of Comparative Labour Law*, 37(1), pp. 65–94. Available at: https://kluwerlawonline.com/journalarticle/International+Journal+of+Comparative+Labour+Law+and+Industrial+Relations/37.1/IJCL2021003 (Accessed: 28 February 2022).

Connelly, R. and Platt, L. (2014) ‘Cohort Profile: UK Millennium Cohort Study (MCS)’, *International Journal of Epidemiology*, 43(6), pp. 1719–1725. Available at: https://doi.org/10.1093/ije/dyu001 (Accessed: 28 February 2022).

COORDINATE Network (2022a) *CORDINATE: Scientific Community*. Available at: https://www.coordinate-network.eu/scientific-community (Accessed: 28 February 2022).

COORDINATE Network (2022b) *COORDINATE: Transnational Access Visits (TAV)*. Available at: https://www.coordinate-network.eu/transnational-visits (Accessed: 28 February 2022).

Corna, L. M. (2013) ‘A life course perspective on socioeconomic inequalities in health: a critical review of conceptual frameworks’, *Advances in Life Course Research*, 18(2), pp. 150–159. Available at: https://doi.org/10.1016/j.alcr.2013.01.002 (Accessed: 28 February 2022).

Cosco, T. D., Kaushal, A., Hardy, R., Richards, M., Kuh, D. and Stafford, M. (2017) ‘Operationalising resilience in longitudinal studies: A systematic review of methodological approaches’, *Journal of Epidemiology and Community Health*, 71(1), pp. 98–104. Available at: https://doi.org/10.1136/jech-2015-206980 (Accessed: 28 February 2022).

Cotter, S., Healy, C., Ni Cathain, D., Williams, P., Clarke, M. and Cannon, M. (2019) ‘Psychopathology and early life stress in migrant youths: An analysis of the ‘Growing Up in Ireland’ study’, *Irish Journal of Psychological Medicine*, 36(3), pp. 177–185. Available at: https://doi:10.1017/ipm.2018.53 (Accessed: 28 February 2022).

Cramer Development Incorporated (2022) *BabySteps—Baby Milestones* [Mobile app]. Available at: App Store. https://apps.apple.com/us/app/babysteps-baby-milestones/id1093215033 (Accessed: 28 February 2022).

CSO (2016a) *Census of population 2016 – Profile 8: Irish Travellers, ethnicity, and religion.* Available at: https://www.cso.ie/en/releasesandpublications/ep/p-cp8iter/p8iter/p8itd/ (Accessed: 28 February 2022).

CSO (2016b) *Census of Population 2016 – Profile 7 Migration and Diversity*. Available at: https://tbinternet.ohchr.org/Treaties/CERD/Shared%20Documents/IRL/INT_CERD_ADR_IRL_32728_E.pdf (Accessed: 28 February 2022).

CSO (2022a) *CSO Data Protocol Register—CSO - Central Statistics Office*. Available at: https://www.cso.ie/en/aboutus/lgdp/csodatapolicies/csodataprotocol/csodataprotocolregister/ (Accessed: 28 February 2022).

CSO (2022b) *Pathfinder Research—CSO - Central Statistics Office*. Available at: https://www.cso.ie/en/methods/research/ (Accessed: 28 February 2022).

CSO. (2018). *Population and Migration Estimates April 2018*. Available at: https://www.cso.ie/en/releasesandpublications/er/pme/populationandmigrationestimatesapril2018/ (Accessed: 28 February 2022).

CSO. (2019). *Information Society Statistics—Households 2019: Household Internet Connectivity*. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-isshh/informationsocietystatistics-households2019/householdinternetconnectivity/ (Accessed: 28 February 2022).

CSO. (2021a). *Census 2022- CSO - Central Statistics Office*. Available at: https://www.cso.ie/en/census/census2022/census2022faq/ (Accessed: 28 February 2022).

CSO. (2021b*). Population and Migration Estimates, April 2021*. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-pme/populationandmigrationestimatesapril2021/ (Accessed: 28 February 2022).

da Silva, S. G., Evenson, K. R., da Silva, I. C. M., Mendes, M. A., Domingues, M. R., da Silveira, M. F., Wehrmeister, F. C., Ekelund, U. and Hallal, P. C. (2018) 'Correlates of accelerometer-assessed physical activity in pregnancy—The 2015 Pelotas (Brazil) Birth Cohort Study', *Scandinavian Journal of Medicine & Science in Sports*, 28(8), pp. 1934–1945. Available at: https://doi.org/10.1111/sms.13083 (Accessed: 28 February 2022).

da Silva, S. G., Evenson, K. R., Ekelund, U., da Silva, I. C. M., Domingues, M. R., da Silva, B. G. C., Mendes, M. de A., Cruz, G. I. N. and Hallal, P. C. (2019) 'How many days are needed to estimate wrist-worn accelerometry-assessed physical activity during the second trimester in pregnancy?', *PLOS ONE*, 14(6), p. e0211442. Available at: https://doi.org/10.1371/journal.pone.0211442 (Accessed: 28 February 2022).

Davies, S., Davenport, L., Lamb, K. and Wake, M. (2020) *Developing GenV's Research Methodologies: Brainstorming Sessions and Survey 2018*. Generation Victoria Working Paper 2020-03. Available at: https://doi:10.25374/MCRI.8234006 (Accessed: 28 February 2022).

Davis-Kean, P., Chamber, R. L., Davidson, L. L., Kleinert, C., Ren, Q. and Tang, S. (2018) *Longitudinal Studies Strategic Review 2017: Report to the Economic and Social Research Council*. Available at: https://esrc.ukri.org/files/news-events-and-publications/publications/longitudinal-studies-strategic-review-2017/ (Accessed: 28 February 2022).

DCYA (2012) *Guidance for developing ethical research projects involving children*. Dublin: Department of Children and Youth Affairs.

DCYA (2017) *Children first: National guidance for the protection and welfare of children*. Dublin: Department of Children and Youth Affairs.

de Gavelle, E., de Lauzon-Guillain, B., Charles, M. A., Chevrier, C., Hulin, M., Sirot, V., Merlo, M., and Nougadère, A (2016) 'Chronic dietary exposure to pesticide residues and associated risk in the French ELFE cohort of pregnant women', *Environment International*, 92, pp. 533–542. Available at: https://doi.org/10.1016/j.envint.2016.04.007 (Accessed: 28 February 2022).

de Groot, S., van der Linden, N., Franken, M. G., Blommestein, H. M., Leeneman, B., van Rooijen, E., Koos van der Hoeven, J. J. M., Wouters, M. W., Westgeest, H. M. and Uyl-de Groot, C. A. (2017) 'Balancing the Optimal and the Feasible: A Practical Guide for Setting Up Patient Registries for the Collection of Real-World Data for Health Care Decision Making Based on Dutch Experiences', *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 20(4), pp. 627–636. Available at: https://doi.org/10.1016/j.jval.2016.02.007 (Accessed: 28 February 2022).

Deave, T., Kendal, S., Lingam, R., Day, C., Goodenough, T., Bailey, E., Ginja, S., Nightingale, S. and Coad, J. (2019) 'A study to evaluate the effectiveness of Best Beginnings' Baby Buddy phone app in England: A protocol paper', *Primary Health Care Research & Development*, 20. Available at: https://doi.org/10.1017/S1463423618000294 (Accessed: 28 February 2022).

Department of Children and Youth Affairs. (2018). *First 5: A whole-of-Government strategy for babies, young children and their families 2019-2028*. Dublin: Government of Ireland. Available at: https://first5.gov.ie/userfiles/pdf/5223_4966_DCYA_EarlyYears_INTERACTIVE_Booklet_280x215_v1.pdf#view=fit (Accessed: 28 February 2022).

Department of Children, Equality, Disability, Integration and Youth (2020a) *An Overview of the Use of Growing Up in Ireland Data Files 2010-19*. Available at: https://www.gov.ie/en/publication/5c5cd-an-overview-of-the-use-of-growing-up-in-ireland-data-files-2010-19/ (Accessed: 28 February 2022).

Department of Children, Equality, Disability, Integration and Youth (2020b) *Using Growing Up in Ireland Data | A Guidance Note for policy makers, researchers and NGOs*. Dublin: Department of Children and Youth Affairs. Available at: https://www.gov.ie/en/publication/b3763-using-growing-up-in-ireland-data/ (Accessed: 28 February 2022).

Department of Health and Children. (2000). *The National Children's Strategy: Our Children: Their Lives* [Report]. Dublin: Stationery Office. Available at: https://www.cypsc.ie/_fileupload/Documents/Resources/National%20Children's%20Strategy%20-%20our%20children,%20their%20lives.pdf (Accessed: 28 February 2022).

Dereumeaux, C., Saoudi, A., Pecheux, M., Berat, B., de Crouy-Chanel, P., Zaros, C., Brunel, S., Delamaire, C., le Tertre, A., Lefranc, A., Vandentorren, S. and Guldner, L. (2016) 'Biomarkers of exposure to environmental contaminants in French pregnant women from the Elfe cohort in 2011', *Environment International*, 97, pp. 56–67. Available at: https://doi.org/10.1016/j.envint.2016.10.013 (Accessed: 28 February 2022).

Devine, D., Symonds, J. E., Sloan, S., Cahoon, A., Crean, M., Farrell, E., Davies, A., Blue, T. and Hogan, J. (2020) *Children's School Lives: An introduction.* Report No 1. Dublin: University College Dublin. Available at: https://a115i2ikt0k3wno971cywnm1-wpengine.netdna-ssl.com/wp-content/uploads/2020/12/CSL-Annual-Report-1-Final.pdf (Accessed: 28 February 2022).

Dex, S. and Joshi, H. (2004) *Millennium Cohort Study First Survey: A User's Guide to Initial Findings.* London: Centre for Longitudinal Studies, Institute of Education, University of London.

Dezateux, C., Colson, D., Brocklehurst, P. and Elias, P. (2016) ''LIFE AFTER LIFE STUDY' Report of a Scientific Meeting held at The Royal College of Physicians, London, UK, 14th January 2016', in *Life Study Working Papers. Life Course Epidemiology and Biostatistics/UCL Institute of Child Health*. London: Life Course Epidemiology and Biostatistics/UCL Institute of Child Health. Available at: http://dx.doi.org/10.14324/000.rp.1485681 (Accessed: 28 February 2022).

DNBC (2020a) *About the DNBC*. Available at: https://www.dnbc.dk/about-the-dnbc (Accessed: 28 February 2022).

DNBC (2020b) *Approvals and Consent: Translation of Information Leaflet given to participants in the Danish National Birth Cohort together with Informed Consent Form. Recruitment period: 1997-2002.* Available at: https://www.dnbc.dk/access-to-dnbc-data/approvals-and-consent (Accessed: 28 February 2022).

DNBC (2020c) *DNBC Data Available*. Available at: https://www.dnbc.dk/data-available (Accessed: 28 February 2022).

Dominick, G. M., Winfree, K. N., Pohlig, R. T. and Papas, M. A. (2016) 'Physical Activity Assessment Between Consumer- and Research-Grade Accelerometers: A Comparative Study in Free-Living Conditions', *JMIR MHealth and UHealth*, 4(3), p. e6281. Available at: https://doi.org/10.2196/mhealth.6281 (Accessed: 28 February 2022).

Downs, J., Setakis, E., Mostafa, T., Hayes, R., Hotopf, M., Ford, T. and Gilbert, R. (2017) 'Linking strategies and biases when matching cohorts to the National Pupil Database', *International Journal of Population Data Science*, 1(1), p.1. Available at: https://doi.org/10.23889/ijpds.v1i1.369 (Accessed: 28 February 2022).

Doyle, A. and Golding, J. (2009) 'The costing and funding of longitudinal birth cohort studies', *Paediatric and Perinatal Epidemiology*, 23, pp. 86–92. Available at: https://doi.org/10.1111/j.1365-3016.2008.01011.x (Accessed: 28 February 2022).

DPC (2019) *Public consultation on the processing of children's personal data and the rights of children as data subjects under the General Data Protection Regulation.* Dublin: Data Protection Commission. Available at: https://www.dataprotection.ie/sites/default/files/uploads/2018-12/DPC_ChildrensRights_2019_English.pdf (Accessed: 28 February 2022).

DPC (2020) *Children Front and Centre: Fundamentals for a Child-Oriented Approach to Data Processing.* Draft Version for Public Consultation. Dublin: Data Protection Commission. Available at: https://www.dataprotection.ie/sites/default/files/uploads/2020-12/Fundamentals%20for%20a%20Child-Oriented%20Approach%20to%20Data%20Processing_Draft%20Version%20for%20Consultation_EN.pdf (Accessed: 28 February 2022).

DPC (2021) *Children Front and Centre. Fundamentals for a Child-Oriented Approach to Data Processing*. Dublin: Data Protection Commission. Available at: https://www.dataprotection.ie/sites/default/files/uploads/2021-12/Fundamentals%20for%20a%20Child-Oriented%20Approach%20to%20Data%20Processing_FINAL_EN.pdf (Accessed: 28 February 2022).

Duşa, A., Nelle, D., Stock, G. and Wagner, G. G. (eds.). (2014). Facing the future: European research infrastructures for the humanities and social sciences. *Facing the Future: European Research Infrastructure for Humanities and Social Sciences* (November 21/22, 2013, Berlin), SCIVERO Verlag. Available at: https://www.allea.org/wp-content/uploads/2015/09/2014_06_04-FACING_THE_FUTURE.pdf (Accessed: 28 February 2022).

eHealth Ireland (2021) *Maternal & Newborn Clinical Management System (MN-CMS)*. Available at: https://www.ehealthireland.ie/strategic-programmes/mncms/https:/www.ehealthireland.ie/strategic-programmes/mncms/maternal-newborn-clinical-management-system.html (Accessed: 28 February 2022).

Elder, G. and Giele, J. Z. (2009) *The Craft of Life Course Research*. New York: Guildford Press.

ELFE (2021) *Family Section*. Available at: https://elfe.site.ined.fr/fr/espace-famille/lettres-information/ (Accessed: 28 February 2022).

ELFE (n.d.a) *The Elfe Study: Key stages*. Available at: https://www.elfe-france.fr/en/the-elfe-study/how-does-it-work/key-stages/ (Accessed: 28 February 2022).

ELFE (n.d.b) *The Elfe Study: Participants*. Available at: https://elfe.site.ined.fr/en/the-elfe-study/how-does-it-work/participants/ (Accessed 14 October 2021).

Etzel, R., Charles, M.-A., Dellarco, M., Gajeski, K., Jöckel, K.-H., Hirschfeld, S., Kamijima, M., Kawamoto, T., Kolossa-Gehring, M., Nakayama, S., Schmidt, B., Tian, Y., Wolz, B., Zaros, C. and Zhang, J. (2014) 'Harmonizing Biomarker Measurements in Longitudinal Studies of Children's Health and the Environment', *Biomonitoring*, 1(1), 50-62.Available at: https://doi.org/10.2478/bimo-2014-0006 (Accessed: 28 February 2022).

Faherty, L. J., Hantsoo, L., Appleby, D., Sammel, M. D., Bennett, I. M. and Wiebe, D. J. (2017) 'Movement patterns in women at risk for perinatal depression: Use of a mood-monitoring mobile application in pregnancy', *Journal of the American Medical Informatics Association: JAMIA*, 24(4), pp. 746–753. Available at: https://doi.org/10.1093/jamia/ocx005 (Accessed: 28 February 2022).

Firestone, R., Cheng, S., Pearce, N., Douwes, J., Merletti, F., Pizzi, C., Pivetta, E., Rusconi, F. and Richiardi, L. (2015) 'Internet-Based Birth-Cohort Studies: Is This the Future for Epidemiology?', *JMIR Research Protocols*, 4(2), p. e71. Available at: https://doi.org/10.2196/resprot.3873 (Accessed: 28 February 2022).

Fischer, F. and Kleen, S. (2021) 'Possibilities, Problems, and Perspectives of Data Collection by Mobile Apps in Longitudinal Epidemiological Studies: Scoping Review', *Journal of Medical Internet Research*, 23(1), p. e17691. Available at: https://doi.org/10.2196/17691 (Accessed: 28 February 2022).

Fitzsimons, E., Goodman, A., Kelly, E. and Smith, J. P. (2017) 'Poverty dynamics and parental mental health: Determinants of childhood mental health in the UK', *Social Science & Medicine*, 175, pp. 43–51. Available at: https://doi.org/10.1016/j.socscimed.2016.12.040 (Accessed: 28 February 2022).

Fitzsimons, E., Moulton, V., Hughes, D. A., Neaves, S., Ho, K., Hemani, G., Timpson, N., Calderwood, L., Gilbert, E. and Ring, S. (2021) 'Collection of genetic data at scale for a nationally representative population: The UK Millennium Cohort Study', *Longitudinal and Life Course Studies*, 13 (1), 169-187. Available at: https://doi.org/10.1332/175795921X16223668101602 (Accessed: 28 February 2022).

Fitzsimons, E., Moulton, V., Neaves, S., Ho, K., Timpson, N., Calderwood, L., Gilbert, E. and Ring, S. (2020) *Collection of DNA samples and genetic data at scale in the UK Millennium Cohort Study*. CLS Working Paper No. 2020/7. London: Centre for Longitudinal Studies. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2020/08/CLS-working-paper-2020-7-Collection-of-DNA-samples-and-genetic-data-at-scale-in-the-UK-Millennium-Cohort-Study.pdf (Accessed: 28 February 2022).

Forrest, C. B., Blackwell, C. K. and Camargo, C. A. (2018) 'Advancing the Science of Children's Positive Health in the NIH Environmental influences on Child Health Outcomes (ECHO) Research Program', *The Journal of Pediatrics*, 196, pp. 298–300. Available at: https://doi.org/10.1016/j.jpeds.2018.02.004 (Accessed: 28 February 2022).

Gares, V., Panico, L., Castagne, R., Delpierre, C., and Kelly-Irving, M. (2017) 'The role of the early social environment on Epstein Barr virus infection: A prospective observational design using the Millennium Cohort Study', *Epidemiology and Infection*, 145(16), 3405–3412. Available at: https://doi.org/10.1017/S0950268817002515 (Accessed: 28 February 2022).

Garritty, C., Gartlehner, G., Nussbaumer-Streit, B., King, V. J., Hamel, C., Kamel, C., Affengruber, L. and Stevens, A. (2021) 'Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews', *Journal of Clinical Epidemiology*, 130, pp. 13–22. Available at: https://doi.org/10.1016/j.jclinepi.2020.10.007 (Accessed: 28 February 2022).

GDPR.EU (2018, November 14). *Recital 32—Conditions for consent*. Available at: https://gdpr.eu/recital-32-conditions-for-consent/ (Accessed: 28 February 2022).

GenV (2022a) *What does being part of GenV involve?* Available at: https://genv.org.au/for-parents/what-does-being-part-of-genv-involve/ (Accessed: 28 February 2022).

GenV (2022b) *Why GenV is significant*. Available at: https://www.genv.org.au/for-policy-makers/why-genv-is-significant/ (Accessed: 28 February 2022).

Gibney, S. and McCarthy, T. (2020) *Research Brief: Profile of Smartphone Ownership and Use in Ireland*. Dublin: Research Services and Policy Unit, Research & Development and Health Analytics Division, Department of Health.

Giesbertz, N. A. A., Bredenoord, A. L. and van Delden, J. J. M. (2016) 'When Children Become Adults: Should Biobanks Re-Contact?', *PLOS Medicine*, 13(2), p. e1001959. Available at: https://doi.org/10.1371/journal.pmed.1001959 (Accessed: 28 February 2022).

Gilbert, E., Conolly, A., Calderwood, L. and Rose, N. (2017) *Measuring young people's physical activity using accelerometers in the UK Millennium Cohort Study*. CLS Working Paper No. 2017/15. London: Centre for Longitudinal Studies. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/11/CLS-WP-201715-Measuring-young-peoples-physical-activity-using-accelerometers-in-the-UK-Millennium-Cohort-Study.pdf (Accessed: 28 February 2022).

Gilbert, R., Lafferty, R., Hagger-Johnson, G., Harron, K., Zhang, L.-C., Smith, P., Dibben, C. and Goldstein, H. (2018) 'GUILD: GUidance for Information about Linking Data sets', *Journal of Public Health*, 40(1), pp. 191–198. Available at: https://doi.org/10.1093/pubmed/fdx037 (Accessed: 28 February 2022).

Glavind, M. T., Forman, A., Arendt, L. H., Nielsen, K. and Henriksen, T. B. (2017) 'Endometriosis and pregnancy complications: A Danish cohort study', *Fertility and Sterility*, 107(1), pp. 160–166. Available at: https://doi.org/10.1016/j.fertnstert.2016.09.020 (Accessed: 28 February 2022).

Gluckman, P. D. and Hanson, M. A. (2004) 'Living with the past: Evolution, development, and patterns of disease', *Science*, 305(5691), pp. 1733–1736. Available at: https://doi.org/10.1126/science.1095292 (Accessed: 28 February 2022).

Gluckman, P. D., Hanson, M. A. and Buklijas, T. (2010) 'A conceptual framework for the developmental origins of health and disease', *Journal of Developmental Origins of Health and Disease*, 1(1), pp. 6–18. Available at: https://doi.org/10.1017/S2040174409990171 (Accessed: 28 February 2022).

Gluckman, P. D., Hanson, M. A., Cooper, C. and Thornburg, K. L. (2008) ‘Effect of In Utero and Early-Life Conditions on Adult Health and Disease’, *The New England Journal of Medicine*, 359(1), pp. 61–73. Available at: https://doi.org/10.1056/NEJMra0708473 (Accessed: 28 February 2022).

Golding, J. (2009) ‘Who should be studied and when in a longitudinal birth cohort?’, *Paediatric and Perinatal Epidemiology*, 23(s1), pp. 15–22. Available at: https://doi.org/10.1111/j.1365-3016.2008.00998.x (Accessed: 28 February 2022).

Golding, J. and Steer, C. (2009) ‘How many subjects are needed in a longitudinal birth cohort study?’, *Paediatric and Perinatal Epidemiology*, 23(s1), pp. 31–38. Available at: https://doi.org/10.1111/j.1365-3016.2008.00997.x (Accessed: 28 February 2022).

Golding, Pembrey, Jones, and Team, T. A. S. (2001) ‘ALSPAC–The Avon Longitudinal Study of Parents and Children’, *Paediatric and Perinatal Epidemiology,* 15(1), 74–87. Available at: https://doi.org/10.1046/j.1365-3016.2001.00325.x (Accessed: 28 February 2022).

Goldstein, E., Bakhireva, L. N., Nervik, K., Hagen, S., Turnquist, A., Zgierska, A. E., Marquez, L. E., McDonald, R., Lo, J. and Chambers, C. (2021) ‘Recruitment and retention of pregnant women in prospective birth cohort studies: A scoping review and content analysis of the literature’, *Neurotoxicology and Teratology*, 85, p. 106974. Available at: https://doi.org/10.1016/j.ntt.2021.106974 (Accessed: 28 February 2022).

Goldstein, H., Sera, F., Elias, P. and Dezateux, C. (2017) ‘Integrating area-based and national samples in birth cohort studies: The case of Life Study’, *Longitudinal and Life Course Studies*, 8(3), pp. 281–289. Available at: https://doi.org/10.14301/llcs.v8i3.439 (Accessed: 28 February 2022).

Goodman, A., Brown, M., Silverwood, R. J., Sakshaug, J. W., Calderwood, L., Williams, J. and Ploubidis, G. (2020) *The impact of using the web in a mixed mode follow-up of a longitudinal birth cohort study: Evidence from the National Child Development Study.* CLS Working Paper No. 2020/9. London: Centre for Longitudinal Studies. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2020/10/CLS-working-paper-2020-9-The-impact-of-using-the-web-in-a-mixed-mode-follow-up-of-a-longitudinal-birth-cohort-study.pdf (Accessed: 28 February 2022).

Gray, M. and Sanson, A. (2005) ‘Growing Up in Australia: The longitudinal study of Australian children’, *Family Matters*, 72, pp. 4–9.

Green, E., Ritchie, F., Webber, D. J., Mytton, J. A., Deave, T., Montgomery, A., Woolfrey, L., ul-Baset, K. and Chowdhury, S. (2015) *Enabling data linkage to maximise the value of public health research data. Final report to The Wellcome Trust.* Public Health Research Data Forum. Available at: https://cms.wellcome.org/sites/default/files/enabling-data-linkage-to-maximise-value-of-public-health-research-data-phrdf-mar15.pdf (Accessed: 28 February 2022).

Greene, N., Greenland, S., Olsen, J. and Nohr, E. A. (2011) 'Estimating Bias From Loss to Follow-up in the Danish National Birth Cohort', *Epidemiology*, 22(6), pp. 815–822. Available at: https://doi.org/10.1097/EDE.0b013e31822939fd (Accessed: 28 February 2022).

Greene, S., Hyland, A., Kelleher, C., Mennell, S., Whelan, B., Wilde, J. and Collins, C. (2001) *Design of the National Children's Strategy Longitudinal Study of Children*. Submitted to the Health Research Board 27/7/2001.

Greene, S., Williams, J., Doyle, E., Harris, E., McCrory, C., Murray, A., Swords, L., Thornton, M., Quail, A., Layte, R., O'Dowd, T. and Whelan, C. (2010b) *Growing Up in Ireland: Review of the literature pertaining to the first wave of data collection with the Infant Cohort at 9 months.* Literature Review Series No. 3. Dublin: Economic and Social Research Institute, Trinity College Dublin, Office of the Minister for Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/BKMNEXT258.pdf (Accessed 26 October 2021).

Greene, S., Williams, J., Layte, R., Doyle, E., Harris, E., McCrory, C., Murray, A., O'Dowd, T., Quail, A., Swords, L., Thornton, M. and Whelan, C. T. (2010a) *Growing Up in Ireland: Background and conceptual framework.* Literature Review Series No. 1. Dublin: Economic and Social Research Institute, Trinity College Dublin, Office of the Minister for Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/BKMNEXT257.pdf (Accessed 26 October 2021).

Griffiths, L. J., Cortina-Borja, M., Sera, F., Pouliou, T., Geraci, M., Rich, C., Cole, T. J., Law, C., Joshi, H., Ness, A. R., Jebb, S. A., and Dezateux, C. (2013) 'How active are our children? Findings from the Millennium Cohort Study', *BMJ Open*, 3(8), p. E002893. Available at: http://dx.doi.org/10.1136/bmjopen-2013-002893 (Accessed: 28 February 2022).

Grobon S, Panico L, Solaz A. (2019) 'Socioeconomic inequalities in language and motor development in 2-year-old children', *Bulletin épidémiologique Hebdomadaire, 1,* 2-9. Available at: http://beh.santepubliquefrance.fr/beh/2019/1/pdf/2019_1_1.pdf (Accessed: 28 February 2022).

Grobon, S., Messayke, L. and Solaz, A. (2019) 'Inégalités socioéconomiques dans le développement langagier et moteur des enfants à 2 ans', *Bulletin épidémiologique hebdomadaire*, 1, pp. 2–9. Available at: http://beh.santepubliquefrance.fr/beh/2019/1/pdf/2019_1_1.pdf (Accessed: 28 February 2022).

Grosser, A., Razum, O., Vrijkotte, T. G., Hinz, I. M. and Spallek, J. (2016) 'Inclusion of migrants and ethnic minorities in European birth cohort studies—a scoping review', *The European Journal of Public Health*, 26(6), pp. 984–991. Available at: https://doi.org/10.1093/eurpub/ckw068 (Accessed: 28 February 2022).

Growing Up in Ireland (2022) *About Growing Up in Ireland.* Available at: https://www.growingup.ie/about-growing-up-in-ireland/ (Accessed: 28 February 2022).

Growing Up in Ireland Study Team (2018) *Key findings: Cohort '08 at 9 years old,* Health and Physical Development No. 3. Dublin: Economic and Social Research Institute, Trinity College Dublin, and Department of Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/Growing-Up-in-Ireland-Health-and-Physical-Development.pdf (Accessed: 28 February 2022).

GUI (2021) *Growing Up in Ireland. Key Findings: Special COVID-19 Survey*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children, Equality, Disability, Integration and Youth. Available at: https://www.growingup.ie/pubs/Covid-KF_Web-ready.pdf (Accessed: 28 February 2022).

GUI (2022) *Cohort '08 at 13 years Participants Information – Growing Up in Ireland.* Available at: https://www.growingup.ie/cohort08at13-participants-information/ (Accessed: 28 February 2022).

Gustavson, K., von Soest, T., Karevold, E. and Røysamb, E. (2012) 'Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study', *BMC Public Health*, 12(918), 1-11. Available at: https://doi.org/10.1186/1471-2458-12-918 (Accessed: 28 February 2022).

Guthmann, J. P., Ragot, M., Boutieb, B., Bois, C., Dufourg, M. N. and Lévy-Bruhl, D. (2016) 'Vaccination coverage and socioeconomic determinants of BCG vaccination in children before 3 months: Results of the Elfe cohort study, 2011', *Revue d'Épidémiologie et de Santé Publique*, 64(4), pp. 271–280. Available at: https://doi.org/1016/j.respe.2016.06.001 (Accessed: 28 February 2022).

Halfon, N., and Forrest, C. B. (2018) 'The Emerging Theoretical Framework of Life Course Health Development', in N. Halfon, C. B. Forrest, R. M. Lerner, & E. M. Faustman (eds.), *Handbook of Life Course Health Development* (pp. 19–43). Springer International Publishing. Available at: https://doi.org/10.1007/978-3-319-47143-3_2 (Accessed: 28 February 2022).

Halfon, N., Larson, K., Lu, M., Tullis, E. and Russ, S. (2014) 'Lifecourse health development: past, present and future', *Maternal and Child Health Journal*, 18(2), pp. 344–365. Available at: https://doi.org/10.1007/s10995-013-1346-2 (Accessed: 28 February 2022).

Hallit, S., Leynaert, B., Delmas, M. C., Rocchi, S., De Blic, J., Marguet, C., Scherer, E., Dufourg, M. N., Bois, C., Reboux, G., Millon, L., Charles, M. A. and Raherison, C. (2018) 'Wheezing phenotypes and risk factors in early life: The ELFE cohort', *PLOS ONE*, 13(4), 1-15. Available at: https://doi.org/10.1371/journal.pone.0196711 (Accessed: 28 February 2022).

Hanafin, S. (2020) *Scoping Review of Literature: Best international practices in linking administrative and longitudinal study data*. Dublin: Department of Children, Equality, Disability, Integration and Youth. Available at: https://igees.gov.ie/wp-content/uploads/2020/10/201029-Scoping-Review-of-Literature.pdf (Accessed: 28 February 2022).

Hansen, K. (2014) *Millennium Cohort Study: A Guide to the Datasets (Eight Edition). First, Second, Third, Fourth and Fifth Surveys*. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-Guide-to-the-Datasets-022014.pdf (Accessed: 28 February 2022).

Hanson, M. A., Cooper, C., Aihie Sayer, A., Eendebak, R. J., Clough, G. F. and Beard, J. R. (2016) 'Developmental aspects of a life course approach to healthy ageing', *The Journal of Physiology*, 594(8), pp. 2147–2160. Available at: https://doi.org/10.1113/JP270579 (Accessed: 28 February 2022).

Hanson, M. and Gluckman, P. (2011) 'Developmental origins of noncommunicable disease: population and public health implications', *The American Journal of Clinical Nutrition*, 94(6), pp. 1754S–1758S. Available at: https://doi.org/10.3945/ajcn.110.001206 (Accessed: 28 February 2022).

Hanson, T., Westwood, P. and Carpenter, H. (2018) *Understanding Society Innovation Panel Wave 11: Technical Report*. Essex: The Institute for Social and Economic Research, University of Essex.

Harrison, S., Alderdice, F., Henderson, J., Redshaw, M. and Quigley, M. A. (2020) 'Trends in response rates and respondent characteristics in five National Maternity Surveys in England during 1995–2018', *Archives of Public Health*, 78(1), p. 46. Available at: https://doi.org/10.1186/s13690-020-00427-w (Accessed: 28 February 2022).

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L. and Goldstein, H. (2017) 'Challenges in administrative data linkage for research', *Big Data & Society*, 4(2), p. 2053951717745678. Available at: https://doi.org/10.1177/2053951717745678 (Accessed: 28 February 2022).

Harron, K., Doidge, J. C. and Goldstein, H. (2020) 'Assessing data linkage quality in cohort studies', *Annals of Human Biology*, 47(2), pp. 218–226. Available at: https://doi.org/10.1080/03014460.2020.1742379 (Accessed: 28 February 2022).

Hens, K., Lévesque, E. and Dierickx, K. (2011) 'Children and biobanks: A review of the ethical and legal discussion', *Human Genetics*, 130(3), pp. 403–413. Available at: https://doi.org/10.1007/s00439-011-1031-8 (Accessed: 28 February 2022).

Heude, B., Forhan, A., Slama, R., Douhaud, L., Bedel, S., Saurel-Cubizolles, M.-J., Hankard, R., Thiebaugeorges, O., De Agostini, M., Annesi-Maesano, I., Kaminski, M., Charles, M.-A., Annesi-Maesano, I., Bernard, J., Botton, J., Charles, M.-A., Dargent-Molina, P., de Lauzon-Guillain, B., Ducimetière, P., … on behalf of the EDEN mother-child cohort study group. (2016) 'Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development', *International Journal of Epidemiology*, 45(2), pp. 353–363. Available at: https://doi.org/10.1093/ije/dyv151 (Accessed: 28 February 2022).

HIQA (2014) *International review of approaches countries have taken to integrate National Health and Social Care Data Collections*. Health Information and Quality Authority, Department of Health. Available at: https://www.hiqa.ie/sites/default/files/2017-01/International-review-National-Health-Social-Care-Data-Collections.pdf (Accessed: 28 February 2022).

HIQA (2021) *Findings from the National Public Engagement on Health Information (2020-2021)*. Health Information and Quality Authority, Department of Health. Available at: https://www.hiqa.ie/sites/default/files/2021-09/Findings-from-the-National-Public-Engagement-on-Health-Information.pdf (Accessed: 28 February 2022).

Hirschfeld, S. (2018) 'Introduction and Goals for the National Children's Study', *Frontiers in Pediatrics*, 5, pp. 1-5. Available at: https://www.frontiersin.org/article/10.3389/fped.2017.00240 (Accessed: 28 February 2022).

Hockley, C., Quigley, M. A., Hughes, G., Calderwood, L., Joshi, H. and Davidson, L. L. (2008) 'Linking Millennium Cohort data to birth registration and hospital episode records', *Paediatric and Perinatal Epidemiology*, 22(1), pp. 99–109. Available at: https://doi.org/10.1111/j.1365-3016.2007.00902.x (Accessed: 28 February 2022).

HSE (2021) *Maternity and Infant Care Scheme*. Available at: https://www.hse.ie/eng/services/list/3/maternity/combinedcare.html (Accessed: 28 February 2022).

Hudak, M. L., Park, C. H., Annett, R. D., Hale, D. E., McGovern, P. M., McLaughlin, T. J., Dole, N., Kaar, J. L. and Balsam, M. J. (2016) 'The National Children's Study: An Introduction and Historical Overview', *Pediatrics*, 137(Supplement_4), pp. S213–S218. Available at: https://doi.org/10.1542/peds.2015-4410B (Accessed: 28 February 2022).

Imenda, S. (2014) 'Is There a Conceptual Difference between Theoretical and Conceptual Frameworks? *Journal of Social Sciences*, 38(2), p. 185.

INED (2020a) *Protocol of the Elfe survey in Maternity Units.* Ined-Inserm-EFS. Available at: https://www.ined.fr/fichier/rte/129/cote-recherche/Acc%C3%A8s%20aux%20donn%C3%A9es%20et%20questionnaires/maternity-survey-protocol.pdf (Accessed: 28 February 2022).

INED (2020b) *Protocol of the Elfe survey at two months*. Ined-Inserm-EFS. Available at: https://www.ined.fr/fichier/rte/129/cote-recherche/Acc%C3%A8s%20aux%20donn%C3%A9es%20et%20questionnaires/2months-survey-protocol_1.pdf (Accessed: 28 February 2022).

INED (2020c) *Protocol of the Elfe survey at one year.* Ined-Inserm-EFS. Available at: https://www.ined.fr/fichier/rte/129/cote-recherche/Acc%C3%A8s%20aux%20donn%C3%A9es%20et%20questionnaires/1year-survey-protocol_1.pdf (Accessed: 28 February 2022).

INED (2021) *French Child Cohort Study (Elfe): Explanatory document for the in-home survey at age 3.5 years*. Ined-Inserm-EFS. Available at: https://www.elfe-france.fr/fichier/rte/178/Cot%C3%A9%20recherche/3years-survey-protocol.pdf (Accessed: 28 February 2022).

IoE (2010) *Impact of the Millennium Cohort Study*. London: Institute of Education, University College London.

Ipsos MORI (2016) *Millennium Cohort Study Sixth Sweep (MCS6): Time Use Diary Documentation.* London: Ipsos MORI Social Research Institute.

ISSDA (2018) *Irish Social Science Data Archive (ISSDA): Growing up In Ireland Cohort '98 (Child Cohort) Wave 3*. Available at: https://www.ucd.ie/issda/data/guichild/guichildwave3/ (Accessed: 28 February 2022).

Jabakhanji, S. B., Pavlova, M., Groot, W., Boland, F. and Biesma, R. (2017) 'Social class variation, the effect of the economic recession and childhood obesity at 3 years of age in Ireland', *The European Journal of Public Health*, 27(2), pp. 234–239. Available at: https://doi.org/10.1093/eurpub/ckw235 (Accessed: 28 February 2022).

Jäckle, A., Beninger, K., Burton, J. and Couper, M. P. (2018) *Understanding data linkage consent in longitudinal surveys*. Understanding Society Working Paper Series No. 2018 – 07. Essex: UK Economic and Social Research Council.

Jäckle, A., Beninger, K., Burton, J. and Couper, M. P. (2021) 'Understanding Data Linkage Consent in Longitudinal Surveys', in P. Lynn (ed.) *Advances in Longitudinal Survey Methodology.* Essex: John Wiley & Sons, pp. 122–148.

Jäckle, A., Gaia, A. and Benzeval, M. (2017) *CLOSER Resource Report: Mixing modes and measurement methods in longitudinal studies*. Essex: CLOSER, University of Essex. Available at: https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-Mixing-modes-and-measurement-methods-in-longitudinal-studies.pdf (Accessed: 28 February 2022).

Jäckle, A., Gaia, A. and Benzeval, M. (2018) *The use of new technologies to measure socioeconomic and environmental concepts in longitudinal studies*. Essex: CLOSER. Available at: https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-The-use-of-new-technology-to-measure-socio-economic-and-environmental-concepts.pdf (Accessed: 28 February 2022).

Jacobsen, T. N., Nohr, E. A. and Frydenberg, M. (2010) 'Selection by socioeconomic factors into the Danish National Birth Cohort', *European Journal of Epidemiology*, 25(5), pp. 349–355. Available at: https://doi.org/10.1007/s10654-010-9448-2 (Accessed: 28 February 2022).

Jaddoe, V. W. V., Felix, J. F., Andersen, A.-M. N., Charles, M.-A., Chatzi, L., Corpeleijn, E., Donner, N., Elhakeem, A., Eriksson, J. G., Foong, R., Grote, V., Haakma, S., Hanson, M., Harris, J. R., Heude, B., Huang, R.-C., Inskip, H., Järvelin, M.-R., Koletzko, B., … LifeCycle Project Group. (2020) 'The LifeCycle Project-EU Child Cohort Network: A federated analysis infrastructure and harmonized data of more than 250,000 children and parents', *European Journal of Epidemiology*, 35(7), pp. 709–724. Available at: https://doi.org/10.1007/s10654-020-00662-z (Accessed: 28 February 2022).

Ji, Y., Song, F., Xu, B., Zhu, Y., Lu, C. and Xia, Y. (2019) 'Association between exposure to particulate matter during pregnancy and birthweight: A systematic review and a meta-analysis of birth cohort studies', *Journal of Biomedical Research*, 33(1), pp. 56–68. Available at: https://doi.org/10.7555/JBR.31.20170038 (Accessed: 28 February 2022).

Johnson, J., Atkinson, A. and Rosenberg, R. (2015) *Millennium Cohort Study: Psychological, developmental and health inventories.* London:Centre for Longitudinal Studies, Institute of Education, University of London.

Johnson, S. and Antill, M. (2011) *Impact Evaluation of the Millennium Cohort Study.* Economic and Social Research Council.

Jones, N. L., Gilman, S. E., Cheng, T. L., Drury, S. S., Hill, C. V. and Geronimus, A. T. (2019) 'Life Course Approaches to the Causes of Health Disparities', *American Journal of Public Health*, 109(Suppl 1), pp. S48–S55. Available at: https://doi.org/10.2105/AJPH.2018.304738 (Accessed: 28 February 2022).

Jones, R. (2009) 'Taking and storing biological samples for longitudinal birth cohorts', *Paediatric and Perinatal Epidemiology*, 23(s1), pp. 114–126. Available at: https://doi.org/10.1111/j.1365-3016.2008.01004.x (Accessed: 28 February 2022).

Joshi, H. and Fitzsimons, E. (2016) 'The UK Millennium Cohort: The making of a multipurpose resource for social science and policy', *Longitudinal and Life Course Studies*, 7(4), pp. 409–430. Available at: https://doi.org/10.14301/llcs.v7i4.410 (Accessed: 28 February 2022).

Juillard, H. (2015) *ELFE Inclusion Survey Carried out in 2011: Weightings at Time 0 (Child Birth)*. The French Institute for Demographic Studies (INED).

Kaye, J., Hurles, M., Griffin, H., Grewal, J., Bobrow, M., Timpson, N., Smee, C., Bolton, P., Durbin, R., Dyke, S., Fitzpatrick, D., Kennedy, K., Kent, A., Muddyman, D., Muntoni, F., Raymond, L. F., Semple, R. and Spector, T. (2014) 'Managing clinically significant findings in research: The UK10K example', *European Journal of Human Genetics*, 22(9), pp. 1100–1104. Available at: https://doi.org/10.1038/ejhg.2013.290 (Accessed: 28 February 2022).

Keane, E., Layte, R., Harrington, J., Kearney, P. M. and Perry, I. J. (2012) 'Measured Parental Weight Status and Familial Socio-Economic Status Correlates with Childhood Overweight and Obesity at Age 9', *PLOS ONE*, 7(8), p. e43503. Available at: https://doi.org/10.1371/journal.pone.0043503 (Accessed: 28 February 2022).

Keilow, M., Sievertsen, H. H., Niclasen, J. and Obel, C. (2019) 'The Strengths and Difficulties Questionnaire and standardized academic tests: Reliability across respondent type and age', *PLOS ONE*, 14(7), p. e0220193. Available at: https://doi.org/10.1371/journal.pone.0220193 (Accessed: 28 February 2022).

Kelfve, S., Kivi, M., Johansson, B. and Lindwall, M. (2020) 'Going web or staying paper? The use of web-surveys among older people', *BMC Medical Research Methodology*, 20(1), p. 252. Available at: https://doi.org/10.1186/s12874-020-01138-0 (Accessed: 28 February 2022).

Kelly, L., Ryan, B., Gallagher, R. and Quail, A. (2021) *A Summary Guide of the COVID-19 Web Survey for Cohorts '08 and '98*. Economic and Social Research Institute, Department of Children, Equality, Disability, Integration and Youth, Central Statistics Office, Trinity College Dublin. Available at: https://www.ucd.ie/issda/t4media/0075-00_Summary%20Guide%20COVID%20Survey.pdf (Accessed: 28 February 2022).

Kelly, M. P., Stewart, E., Morgan, A., Killoran, A., Fischer, A., Threlfall, A. and Bonnefoy, J. (2009) 'A conceptual framework for public health: NICE's emerging approach', *Public Health*, 123(1), pp. e14–e20. Available at: https://doi.org/10.1016/j.puhe.2008.10.031 (Accessed: 28 February 2022).

Kelly, Y., Becares, L. and Nazroo, J. (2013) 'Associations between maternal experiences of racism and early child health and development: findings from the UK Millennium Cohort Study', *Journal of Epidemiology and Community Health*, 67, pp. 35–41. Available at: https://dx.doi.org/10.1136/jech-2011-200814 (Accessed: 28 February 2022).

Kent, G. and Pitsia, V. (2018) 'A comparison of the home learning environment of families at risk of socio-economic disadvantage to national norms in Ireland', *Irish Educational Studies*, 37(4), pp. 505–521, Available at: https://doi.org/10.1080/03323315.2018.1471409 (Accessed: 28 February 2022).

Keogh, B., Brady, A. M., Downes, C., Doyle, L., Higgins, A. and McCann, T. (2020) 'Evaluation of a Traveller Mental Health Liaison Nurse: Service User Perspectives', *Issues in Mental Health Nursing*, 41(9), pp. 799–806. Available at: https://doi.org/10.1080/01612840.2020.1731889 (Accessed: 28 February 2022).

Ketende, S. (2010) *Millennium Cohort Study: Technical Report on Response (Third Edition)*. London: Centre for Longitudinal Studies. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/mcs_technical_report_on_response_thrid_edition.pdf (Accessed: 28 February 2022).

Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P. and Kreuter, F. (2019) 'Willingness to Participate in Passive Mobile Data Collection', *Public Opinion Quarterly*, 83(S1), pp. 210–235. Available at: https://doi.org/10.1093/poq/nfz007 (Accessed: 28 February 2022).

Kiernan, K. E. and Huerta, M. C. (2008) ‘Economic deprivation, maternal depression, parenting and children’s cognitive and emotional development in early childhood’, *The British Journal of Sociology*, 59, pp. 783–806. Available at: https://doi.org/10.1111/j.1468-4446.2008.00219.x (Accessed: 28 February 2022).

Knight, J. A., Comino, E. J., Harris, E. and Jackson-Pulver, L. (2009) ‘Indigenous Research: A Commitment to Walking the Talk. The Gudaga Study—an Australian Case Study’, *Journal of Bioethical Inquiry*, 6(4), p. 467. Available at: https://doi.org/10.1007/s11673-009-9186-x (Accessed: 28 February 2022).

Kooijman, M. N., Kruithof, C. J., van Duijn, C. M., Duijts, L., Franco, O. H., van Ijzendoorn, M. H., de Jongste, J. C., Klaver, C. C. W., van der Lugt, A., Mackenbach, J. P., Moll, H. A., Peeters, R. P., Raat, H., Rings, E. H. H. M., Rivadeneira, F., van der Schroeff, M. P., Steegers, E. A. P., Tiemeier, H., Uitterlinden, A. G., … and Jaddoe, V. W. V. (2016) ‘The Generation R Study: Design and cohort update 2017’, *European Journal of Epidemiology*, 31(12), pp. 1243–1264. Available at: https://doi.org/10.1007/s10654-016-0224-9 (Accessed: 28 February 2022).

Kuh, D. and Ben-Shlomo, Y. (2004) *A life course approach to chronic disease epidemiology*. Oxford: Oxford University Press.

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J. and Power, C. (2003) ‘Life course epidemiology’, *Journal of Epidemiology and Community Health*, 57(10), pp. 778–783. Available at: https://doi.org/10.1136/jech.57.10.778 (Accessed: 28 February 2022).

Kyvik, S. and Reymert, I. (2017) ‘Research collaboration in groups and networks: differences across academic fields’, *Scientometrics*, 113, pp. 951–967. Available at: https://doi.org/10.1007/s11192-017-2497-5 (Accessed: 28 February 2022).

Larsen, P. S., Kamper-Jørgensen, M., Adamson, A., Barros, H., Bonde, J. P., Brescianini, S., Brophy, S., Casas, M., Devereux, G., Eggesbø, M., Fantini, M. P., Frey, U., Gehring, U., Grazuleviciene, R., Henriksen, T. B., Hertz-Picciotto, I., Heude, B., Hryhorczuk, D. O., Inskip, H., … and Andersen, A.-M. N. (2013) ‘Pregnancy and Birth Cohort Resources in Europe: A Large Opportunity for Aetiological Child Health Research’, *Paediatric and Perinatal Epidemiology*, 27(4), pp. 393–414. Available at: https://doi.org/10.1111/ppe.12060 (Accessed: 28 February 2022).

Lawlor, D. A. and Mishra, G. D. (2009) *Family Matters: Designing, Analysing and Understanding Family Based Studies in Life Course Epidemiology*. Oxford: Oxford University Press.

Lawlor, D. A., Andersen, A.-M. N. and Batty, G. D. (2009) ‘Birth cohort studies: Past, present and future’, *International Journal of Epidemiology*, 38(4), pp. 897–902. Available at: https://doi.org/10.1093/ije/dyp240 (Accessed: 28 February 2022).

Layte, R. and McCrory, C. (2014) *Growing up in Ireland: Maternal Health Behaviours and Child Growth in Infancy*. Infant Cohort Research Report No. 4.. Dublin: Stationery Office.

Legewie, N. and Nassauer, A. (2018) 'YouTube, Google, Facebook: 21st Century Online Video Research and Research Ethics', *Forum: Qualitative Social Research*, 19(3), 1-21. Available at: https://doi.org/10.17169/fqs-19.3.3130 (Accessed: 28 February 2022).

Leone Sciabolazza, V., Vacca, R., Kennelly Okraku, T. and McCarty, C. (2017) 'Detecting and analyzing research communities in longitudinal scientific networks', *PLOS ONE*, 12(8), p. e0182516. Available at: https://doi.org/10.1371/journal.pone.0182516 (Accessed: 28 February 2022).

Leung, B. M., McDonald, S. W., Kaplan, B. J., Giesbrecht, G. F. and Tough, S. C. (2013) 'Comparison of sample characteristics in two pregnancy cohorts: Community-based versus population-based recruitment methods', *BMC Medical Research Methodology*, 13(1), p. 149. Available at: https://doi.org/10.1186/1471-2288-13-149 (Accessed: 28 February 2022).

Liew, Z., Ritz, B., Virk, J. and Olsen, J. (2016) 'Maternal use of acetaminophen during pregnancy and risk of autism spectrum disorders in childhood: A Danish National Birth Cohort study', *Autism Research: Official Journal of the International Society for Autism Research*, 9(9), pp. 951–958. Available at: https://doi.org/10.1002/aur.1591 (Accessed: 28 February 2022).

Liu, T., Lingam, R., Lycett, K., Mensah, F. K., Muller, J., Hiscock, H., Huque, M. H. and Wake, M. (2018) 'Parent-reported prevalence and persistence of 19 common child health conditions', *Archives of Disease in Childhood*, 103(6), pp. 548–556. Available at: https://doi.org/10.1136/archdischild-2017-313191 (Accessed: 28 February 2022).

Londra, M., Calderwood, L. and Millenium Cohort Team (2006) *Millennium Cohort Study Second Survey: CAPI Questionnaire Documentation*. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS2_CAPI_Questionnaire_Documentation_June_2006_v1-2.pdf (Accessed: 28 February 2022).

Loxton, D., Harris, M. L., Forder, P., Powers, J., Townsend, N., Byles, J. and Mishra, G. (2019) 'Factors Influencing Web-Based Survey Response for a Longitudinal Cohort of Young Women Born Between 1989 and 1995', *Journal of Medical Internet Research*, 21(3), p. e11286. Available at: https://doi.org/10.2196/11286 (Accessed: 28 February 2022).

Lynch, J. and Smith, G. D. (2005) 'A life course approach to chronic disease epidemiology', *Annual Review of Public Health*, 26, pp. 1–35. Available at: https://doi.org/10.1146/annurev.publhealth.26.021304.144505 (Accessed: 28 February 2022).

Lynn, P. (2019) *Sample design: Criteria and procedures*. EuroCohort Work Package No. 8. Available at: https://www.eurocohort.eu/files/filemanager/files/D8.1.pdf (Accessed: 28 February 2022).

Madsen, M. T., Bjerregaard, A., Furtado, J. D., Halldorsson, T. I., Ström, M., Granström, C., Giovannucci, E. and Olsen, S. F. (2019) 'Comparisons of Estimated Intakes and Plasma Concentrations of Selected Fatty Acids in Pregnancy', *Nutrients*, 11(3), p. 568. Available at: https://doi.org/10.3390/nu11030568 (Accessed: 28 February 2022).

Maggi, S., Irwin, L. J., Siddiqi, A. and Hertzman, C. (2010) 'The social determinants of early child development: An overview', *Journal of Paediatrics and Child Health*, 46(11), pp. 627–635. Available at: https://doi.org/10.1111/j.1440-1754.2010.01817.x (Accessed: 28 February 2022).

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjaerven, R., Stoltenberg, C. and MoBa Study Group (2006) 'Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa)', *International Journal of Epidemiology*, 35(5), pp. 1146–1150. Available at: https://doi.org/10.1093/ije/dyl170 (Accessed: 28 February 2022).

Manca, D. P., O'Beirne, M., Lightbody, T., Johnston, D. W., Dymianiw, D.-L., Nastalska, K., Anis, L., Loehr, S., Gilbert, A. and Kaplan, B. J. (2013) 'The most effective strategy for recruiting a pregnancy cohort: A tale of two cities', *BMC Pregnancy and Childbirth*, 13(75), pp. 1-7. Available at: https://doi.org/10.1186/1471-2393-13-75 (Accessed: 28 February 2022).

Marmot, M. and Wilkinson, R. (eds.) (2005) *Social determinants of health*. Oxford: Oxford University Press.

Martínez-Mesa, J., González-Chica, D. A., Duquia, R. P., Bonamigo, R. R. and Bastos, J. L. (2016) 'Sampling: How to select participants in my research study?', *Anais Brasileiros de Dermatologia*, 91(3), pp. 326–330. Available at: https://doi.org/10.1590/abd1806-4841.20165254 (Accessed: 28 February 2022).

Mazzonna, F. (2014) 'The long-lasting effects of family background: A European cross-country comparison', *Economics of Education Review*, 40, pp. 25–42. Available at: https://doi.org/10.1016/j.econedurev.2013.11.010 (Accessed: 28 February 2022).

McCarthy, M. (2014) 'US cancels plan to study 100 000 children from "womb" to age 21', *BMJ*, 349, p. g7775. Available at: https://doi.org/10.1136/bmj.g7775 (Accessed: 28 February 2022).

McCrory, C. and Layte, R. (2012) 'Breastfeeding and risk of overweight and obesity at nine-years of age', *Social Science & Medicine*, 75(2), pp. 323–330. Available at: https://doi.org/10.1016/j.socscimed.2012.02.048 (Accessed: 28 February 2022).

McCrory, C. and McNally, S. (2013) 'The effect of pregnancy intention on maternal prenatal behaviours and parent and child health: results of an Irish cohort study', *Paediatric and Perinatal Epidemiology*, 27(2), pp. 208–215. Available at: https://doi.org/10.1111/ppe.12027 (Accessed: 28 February 2022).

McCrory, C., Leahy, S., Ribeiro, A. I., Fraga, S., Barros, H., Avendano, M., Vineis, P., Layte, R., for the LIFEPATH consortium (2019) 'Maternal educational inequalities in measured body mass index trajectories in three European countries', *Paediatric and Perinatal Epidemiology*, 33(3), pp. 226–237. Available at: https://doi.org/10.1111/ppe.12552 (Accessed: 28 February 2022).

McCrory, C., O'Leary, N., Fraga, S., Ribeiro, A. I., Barros, H., Kartiosuo, N., Raitakari, O., Kivimäki, M., Vineis, P., and Layte, R. (2017) 'Socioeconomic differences in children's growth trajectories from infancy to early adulthood: evidence from four European countries', *Journal of Epidemiology and Community Health*, 71(10), pp. 981–989. Available at: http://dx.doi.org/10.1136/jech-2016-208556 (Accessed: 28 February 2022).

McCrory, C., Williams, J., Murray, A., Quail, A. and Thornton, M. (2013) *Growing Up in Ireland: Design, instrumentation and procedures for the Infant Cohort at wave two (3 years)*. Infant Technical Report No. 3. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

McDade, T. W. and Harris, K. M. (2018) 'The Biosocial Approach to Human Development, Behavior, and Health Across the Life Course', *The Russell Sage Foundation Journal of the Social Sciences: RSF*, 4(4), pp. 2–26. Available at: https://doi.org/10.7758/RSF.2018.4.4.01 (Accessed: 28 February 2022).

McDonald, S. W., Lyon, A. W., Benzies, K. M., McNeil, D. A., Lye, S. J., Dolan, S. M., Pennell, C. E., Bocking, A. D. and Tough, S. C. (2013) 'The All Our Babies pregnancy cohort: Design, methods, and participant characteristics', *BMC Pregnancy and Childbirth*, 13(Suppl 1), p. S2. Available at: https://doi.org/10.1186/1471-2393-13-S1-S2 (Accessed: 28 February 2022).

McGinnity, F., Grotti, R., Russell, H. and Fahey, E. (2018) *Attitudes to diversity in Ireland: March 2018*. Dublin: Economic and Social Research Institute.

McGinnity, F., McMullin, P., Murray, A. and Russell, H. (2017) 'Social inequality in cognitive outcomes in Ireland: What is the role of the home learning environment and childcare?', in Blossfeld, H. P., Kulic, N., Skopek, J. and Triventi, M. (eds.) *Childcare, Early Education and Social Inequality: An International Perspective*. Cheltenham: Edward Elgar Publishing, pp. 109–132. Available at: https://www.esri.ie/publications/social-inequality-in-cognitive-outcomes-in-ireland-what-is-the-role-of-the-home-learning-environment-and-childcare (Accessed: 28 February 2022).

McGinnity, F., Murray, A. and McNally, S. (2013) *Growing Up in Ireland: National Longitudinal Study of Children: Mothers' return to work and childcare choices for infants in Ireland. Infant Cohort Research Report No. 2. Dublin:* Department of Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/BKMNEXT238.pdf (Accessed: 28 February 2022).

McNamara, E., Murphy, D., Murray, A., Smyth, E. and Watson, D. (2020) *Growing Up in Ireland: The lives of 17/18-Year-olds of Cohort'98 (Child Cohort).* Dublin: Economic and Social Research Institute. Available at: https://www.esri.ie/publications/growing-up-in-ireland-the-lives-of-17-18-year-olds-of-cohort-98-childcohort (Accessed: 28 February 2022).

McNamara, E., Murray, A. and Williams, J. (2019) *Growing Up in Ireland: Design, instrumentation and procedures (including summary literature review, pilot report and findings) for Cohort '08 at wave four (7/8 years)*. Technical Series Report Number 2019-3. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

McNamara, E., Murray, A., O'Mahony, D., O'Reilly, C., Smyth, E. and Watson, D. (2021) *Growing Up in Ireland: The Lives of 9-Year-Olds of Cohort '08.* Report No. 10. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children, Equality, Disability, Integration and Youth.

Melchior, M., Chollet, A., Glangeaud-Freudenthal, N., Saurel-Cubizolles, M. J., Dufourg, M. N., Van Der Waerden, J. and Sutter-Dallay, A. L. (2015) 'Tobacco and alcohol use in pregnancy in France: the role of migrant status: the nationally representative ELFE study', *Addictive Behaviours*, 51, pp. 65–71. Available at: https://doi.org/10.1016/j.addbeh.2015.07.015 (Accessed: 28 February 2022).

Messayke, S., Franco, P., Forhan, A., Dufourg, M. N., Charles, M. A. and Plancoulaine, S. (2020) 'Sleep habits and sleep characteristics at age one year in the ELFE birth cohort study', *Sleep Medicine*, 67, pp. 200–206. Available at: https://doi.org/10.1016/j.sleep.2019.11.1255 (Accessed: 28 February 2022).

Mitku, A. A., Zewotir, T., North, D., Jeena, P. and Naidoo, R. N. (2021) 'Effects of prenatal exposure factors on birth outcomes through mediation of favorable fetal growth conditions using structural equation modeling', *PLOS ONE*, 16(4), p. e0249664. Available at: https://doi.org/10.1371/journal.pone.0249664 (Accessed: 28 February 2022).

Mohadjer, L., Krenzke, T., de Kerckhove, W. V. and Li, L. (2016) 'Sampling Design', in *Technical report of the Survey of Adult Skills (PIAAC) (2nd edition)*. Organisation for Economic Co-operation and Development, pp. 1- 40.

Mohan, G. (2021) 'Young, poor, and sick: The public health threat of energy poverty for children in Ireland', *Energy Research & Social Science*, 71, p. 101822. Available at: https://doi.org/10.1016/j.erss.2020.101822 (Accessed: 28 February 2022).

Monnier, M., Moulin, F., Thierry, X., Vandentorren, S., Côté, S., Barbosa, S., Falissard, B., Plancoulaine, S., Charles, M.-A., Simeon, T., Geay, B., Marchand, L., Ancel, P.-Y., Melchior, M., Rouquette, A., Bajos, N., Carrat, F., Ancel, P.-Y., Charles, M.-A., … the SAPRIS study group. (2021) 'Children's mental and behavioral health, schooling, and socioeconomic characteristics during school closure in France due to COVID-19: the SAPRIS project', *Scientific Reports*, 11, pp. 22373. Available at: https://doi.org/10.1038/s41598-021-01676-7 (Accessed: 28 February 2022).

Morales-Suárez-Varela, M., Nohr, E. A., Olsen, J. and Bech, B. H. (2018) 'Potential combined effects of maternal smoking and coffee intake on foetal death within the Danish National Birth Cohort', *European Journal of Public Health*, 28(2), pp. 315–320. Available at: https://doi.org/10.1093/eurpub/ckx222 (Accessed: 28 February 2022).

Morton, S.M.B., Atatoa Carr, P.E., Grant, C.C., Berry, S.D., Bandara, D.K., Mohal, J., Tricker, P. J., Ivory, V.C., Kingi, T.R., Liang, R., Perese, L.M., Peterson, E., Pryor, J.E., Reese, E., Waldie, K.E., and Wall, C.R. (2014) *Growing Up in New Zealand: a longitudinal study of New Zealand children and their families. Now we are Two: Describing our first 1000 days*. Auckland: Growing Up in New Zealand. Available at: https://cdn.auckland.ac.nz/assets/growingup/research-findings-impact/GUiNZ_Now%20we%20are%20four%20report.pdf (Accessed: 28 February 2022).

Morton, S. M. B., Atatoa Carr, P. E., Grant, C. C., Robinson, E. M., Bandara, D. K., Bird, A., Ivory, V. C., Kingi, T. K. R., Liang, R., Marks, E. J., Perese, L. M., Peterson, E. R., Pryor, J. E., Reese, E., Schmidt, J. M., Waldie, K. E. and Wall, C. (2013) 'Cohort profile: Growing Up in New Zealand', *International Journal of Epidemiology*, 42(1), pp. 65–75. Available at: https://doi.org/10.1093/ije/dyr206 (Accessed: 28 February 2022).

Murray, A., McNamara, E., O'Mahony, D., Smyth, E. and Watson, D. (2021) *Growing Up in Ireland: Key findings from the special COVID-19 survey of Cohorts '98 and '08*. Dublin: Economic and Social Research Institute. Available at: https://www.esri.ie/publications/growing-up-in-ireland-key-findings-from-the-special-covid-19-survey-of-cohorts-98-and (Accessed: 28 February 2022).

Murray, A., McNamara, E., Thornton, M., Williams, J. and Smyth, E. (2019) *Growing Up in Ireland – The lives of 5-year-olds*. Dublin: Department of Children and Youth Affairs, Economic and Social Research Institute, and Trinity College Dublin. Available at: https://www.growingup.ie/pubs/Growing-Up-in-Ireland-The-Lives-of-5-Year-Olds.pdf (Accessed: 28 February 2022).

Murray, A., Quail, A., McCrory, C. and Williams, J. (2013) *A Summary Guide to Wave 2 of the Infant Cohort (3 Years) of Growing Up in Ireland*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

Murray, A., Smyth, E., Kelly, L., O'Reilly, C., MacNamara, E. and O'Mahony, D. (2020) *Growing Up in Ireland: Review of selected literature pertaining to the Growing Up in Ireland Cohort '08 (Infant Cohort) at age 9 years*. Literature Review Series No. 2020-4. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/Cohort-08-at-9-literature-review-2020.pdf (Accessed: 28 February 2022).

National Centre for Social Research (2003) *Millennium Cohort Study First Survey: CAPI Questionnaire Documentation*. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://cls.ucl.ac.uk/wpcontent/uploads/2017/07/MCS1_CAPI_Questionnaire_Documentation_March_2006_v1.1.pdf (Accessed: 28 February 2022).

National Economic & Social Council. (2021). *Digital Inclusion in Ireland: Connectivity, Devices & Skills*. NESC Report. (No. 154). Dublin: National Economic & Social Development Office NESDO. http://files.nesc.ie/nesc_reports/en/154_Digital.pdf (Accessed: 28 February 2022).

NCS (2016) *The National Children's Study archive: Study description and guide*. The National Children's Study. Available at: https://www.nichd.nih.gov/sites/default/files/research/NCS/Documents/NCS_Archive_Study_Description.pdf (Accessed: 28 February 2022).

Neuman, Å., Hohmann, C., Orsini, N., Pershagen, G., Eller, E., Kjaer, H. F., Gehring, U., Granell, R., Henderson, J., Heinrich, J., Lau, S., Nieuwenhuijsen, M., Sunyer, J., Tischer, C., Torrent, M., Wahn, U., Wijga, A. H., Wickman, M., Keil, T., and Bergström, A. (2012) 'Maternal smoking in pregnancy and asthma in preschool children: A pooled analysis of eight birth cohorts', *American Journal of Respiratory and Critical Care Medicine*, 186(10), pp. 1037–1043. Available at: https://doi.org/10.1164/rccm.201203-0501OC (Accessed: 28 February 2022).

New Zealand Treasury (2020) *Our living standards framework*. Available at: https://www.treasury.govt.nz/information-and-services/nz-economy/higher-living-standards/our-living-standards-framework (Accessed: 28 February 2022).

Niemelä, M. S., Kangas, M., Ahola, R. J., Auvinen, J. P., Leinonen, A.-M., Tammelin, T. H., Vaaramo, E. S., Keinänen-Kiukaanniemi, S. M., Korpelainen, R. I. and Jämsä, T. J. (2019) 'Dose-response relation of self-reported and accelerometer-measured physical activity to perceived health in middle age—The Northern Finland Birth Cohort 1966 Study', *BMC Public Health*, 19(1), p. 21. Available at: https://doi.org/10.1186/s12889-018-6359-8 (Accessed: 28 February 2022).

Nkyekyer, J., Clifford, S. A., Mensah, F. K., Wang, Y., Chiu, L. and Wake, M. (2021) 'Maximizing Participant Engagement, Participation, and Retention in Cohort Studies Using Digital Methods: Rapid Review to Inform the Next Generation of Very Large Birth Cohorts', *Journal of Medical Internet Research*, 23(5), p. e23499. Available at: https://doi.org/10.2196/23499 (Accessed: 28 February 2022).

Nohr, E. A. and Liew, Z. (2018) 'How to investigate and adjust for selection bias in cohort studies', *Acta Obstetricia et Gynecologica Scandinavica*, 97(4), pp. 407–416. Available at: https://doi.org/10.1111/aogs.13319 (Accessed: 28 February 2022).

Nohr, E. A., Frydenberg, M., Henriksen, T. B. and Olsen, J. (2006) 'Does low participation in cohort studies induce bias?', *Epidemiology*, 17(4), pp. 413–418. Available at: https://doi.org/10.1097/01.ede.0000220549.14177.60 (Accessed: 28 February 2022).

Nolan, A. and Layte, R. (2017) *Understanding use of general practitioner services among children in Ireland*. Dublin: The Stationery Office. Available at: https://www.esri.ie/publications/understanding-use-of-general-practitioner-services-among-children-in-ireland (Accessed: 28 February 2022).

Nolan, A. and Smyth, E. (2021) *Health and Wellbeing in Childhood and Adolescence*. ESRI Research Bulletin 202127, Dublin: ESRI. Available at: https://www.esri.ie/publications/health-and-wellbeing-in-childhood-and-adolescence (Accessed: 28 February 2022).

NREC (2021) *NREC guidance on data protection for research purposes for applicants*. Dublin: National Office for Research Ethics Committees.

Nybo Andersen, A.-M. and Olsen, J. (2011) 'The Danish National Birth Cohort: Selected scientific contributions within perinatal epidemiology and future perspectives', *Scandinavian Journal of Public Health*, 39(7), pp. 115–120. Available at: https://doi.org/10.1177/1403494811407674 (Accessed: 28 February 2022).

O'Donovan, S. M., Murray, D. M., Hourihane, J. O., Kenny, L. C., Irvine, A. D. and Kiely, M. (2015) 'Cohort profile: The Cork BASELINE Birth Cohort Study: Babies after SCOPE: Evaluating the Longitudinal Impact on Neurological and Nutritional Endpoints', *International Journal of Epidemiology*, 44(3), pp. 764–775. Available at: https://doi.org/10.1093/ije/dyu157 (Accessed: 28 February 2022).

Olsen, J. (2012) 'Nine months that last a lifetime. Experience from the Danish National Birth Cohort and lessons learned', *International Journal of Hygiene and Environmental Health*, 215(2), pp. 142–144. Available at: https://doi.org/10.1016/j.ijheh.2011.10.015 (Accessed: 28 February 2022).

Olsen, J. and Meder, K. (2014) 'Better health for mother and child – The Danish National Birth Cohort (DNBC), its structure, history and aims', *Norsk Epidemiologi*, 24(1-2)), pp. 37-38. Available at: https://doi.org/10.5324/nje.v24i1-2.1756 (Accessed: 28 February 2022).

Olsen, J., Melbye, M., Olsen, S. F., Sørensen, T. I., Aaby, P., Andersen, A. M., Taxbøl, D., Hansen, K. D., Juhl, M., Schow, T. B., Sørensen, H. T., Andresen, J., Mortensen, E. L., Olesen, A. W. and Søndergaard, C. (2001) 'The Danish National Birth Cohort--its background, structure and aim', *Scandinavian Journal of Public Health*, 29(4), pp. 300–307. Available at: https://doi.org/10.1177/14034948010290040201 (Accessed: 28 February 2022).

Olsen, S. F., Birgisdottir, B. E., Halldorsson, T. I., Brantsæter, A. L., Haugen, M., Torjusen, H., Petersen, S. B., Strøm, M. and Meltzer, H. M. (2014) 'Possibilities and considerations when merging dietary data from the world's two largest pregnancy cohorts: The Danish National Birth Cohort and the Norwegian Mother and Child Cohort Study', *Acta Obstetricia et Gynecologica Scandinavica*, 93(11), pp. 1131–1140. Available at: https://doi.org/10.1111/aogs.12513 (Accessed: 28 February 2022).

Olsen, S. F., Mikkelsen, T. B., Knudsen, V. K., Orozova-Bekkevold, I., Halldórsson, T. I., Strøm, M. and Østerdal, M. L. (2007) 'Data collected on maternal dietary exposures in the Danish National Birth Cohort', *Paediatric and Perinatal Epidemiology*, 21(1), pp. 76–86. Available at: https://doi.org/10.1111/j.1365-3016.2007.00777.x (Accessed: 28 February 2022).

Orri, M., Russell, A. E., Mars, B., Turecki, G., Gunnell, D., Heron, J., Tremblay, R. E., Boivin, M., Nuyt, A.-M., Côté, S. M. and Geoffroy, M.-C. (2020) 'Perinatal adversity profiles and suicide attempt in adolescence and young adulthood: Longitudinal analyses from two 20-year birth cohort studies', *Psychological Medicine*, 52(7), pp. 1255 - 1267. Available at: https://doi.org/10.1017/S0033291720002974 (Accessed: 28 February 2022).

Pailhé, A., Solaz, A. and Tô, M. (2018) *Can daddies learn how to change nappies? Evidence from a short paternity leave policy*. Paris: INED. Available at: https://hal.archives-ouvertes.fr/hal-02079059/document (Accessed: 28 February 2022).

Panico, L., Tô, M. and Thévenon, O. (2015) 'La fréquence des naissances de petit poids: quelle influence a le niveau d'instruction des mères?', *Population & Sociétés*, 523, pp. 1–4. Available at: https://doi.org/10.3917/popsoc.523.0001 (Accessed: 28 February 2022).

Pansieri, C., Pandolfini, C., Clavenna, A., Choonara, I. and Bonati, M. (2020) 'An Inventory of European Birth Cohorts', *International Journal of Environmental Research and Public Health*, 17(9), p. 3071. Available at: https://doi.org/10.3390/ijerph17093071 (Accessed: 28 February 2022).

Parsons, S. J. and Platt, L. (2016) 'Enhancing Social Surveys Through the Postal Collection of Shed Milk Teeth: An Example of a Large-Scale Cost-Effective Collection on a Longitudinal Study', *Survey Research Methods*, 10, pp. 1–14. Available at: https://doi.org/10.18148/SRM/2016.V10I1.6233 (Accessed: 28 February 2022).

Patalay, P., and Fitzsimons, E. (2021) 'Psychological distress, self-harm and attempted suicide in UK 17-year olds: Prevalence and sociodemographic inequalities', *The British Journal of Psychiatry*, 219(2), 437–439. Available at: https://doi.org/10.1192/bjp.2020.258 (Accessed: 28 February 2022).

Pearson, H. (2015) 'Massive UK baby study cancelled', *Nature*, 526(7575), pp. 620–621. Available at: https://doi.org/10.1038/526620a (Accessed: 28 February 2022).

Peycheva, D., Ploubidis, G. and Calderwood, L. (2021) 'Determinants of consent to administrative records linkage in longitudinal surveys: Evidence from next steps', in in P. Lynn (ed.) *Advances in Longitudinal Survey Methodology*. Essex: John Wiley & Sons.

Piler, P., Kandrnal, V. and Bláha, L. (2017) 'Critical assessment of the research outcomes of European birth cohorts: Linking environmental factors with non-communicable diseases', *Public Health*, 145, pp. 136–145. Available at: https://doi.org/10.1016/j.puhe.2016.12.037 (Accessed: 28 February 2022).

Pirus, C., Bois, C., Dufourg, M., Lanoë, J., Vandentorren, S., Leridon, H. and Elfe, L. (2010) 'Constructing a Cohort: Experience with the French Elfe Project', *Population*, 65, pp. 543–574. Available at: https://doi.org/10.3917/popu.1004.0637 (Accessed: 28 February 2022).

Pirus, C., Leridon, H. and Wiles-Portier, E. (2010) 'Large Child Cohort Studies across the World', *Population*, 65, pp. 575–629. Available at: https://doi.org/10.3917/popu.1004.0671 (Accessed: 28 February 2022).

Polańska, K., Hanke, W., Jurewicz, J., Sobala, W., Madsen, C., Nafstad, P. and Magnus, P. (2011) 'Polish mother and child cohort study (REPRO_PL)—Methodology of follow-up of the children', *International Journal of Occupational Medicine and Environmental Health*, 24(4), pp. 391–398. Available at: https://doi.org/10.2478/s13382-011-0026-y (Accessed: 28 February 2022).

Poulain, T., Baber, R., Vogel, M., Pietzner, D., Kirsten, T., Jurkutat, A., Hiemisch, A., Hilbert, A., Kratzsch, J., Thiery, J., Fuchs, M., Hirsch, C., Rauscher, F. G., Loeffler, M., Körner, A., Nüchter, M., Kiess, W. and LIFE Child study team (2017) 'The LIFE Child study: a population-based perinatal and pediatric cohort in Germany', *European Journal of Epidemiology*, 32(2), pp. 145–158. Available at: https://doi.org/10.1007/s10654-016-0216-9 (Accessed: 28 February 2022).

Power, C., Kuh, D. and Morton, S. (2013) 'From developmental origins of adult disease to life course research on adult disease and aging: Insights from birth cohort studies', *Annual Review of Public Health*, 34(1), pp. 7–28.

Pugh, C. A., Summers, K. M., Bronsvoort, B. M. C., Handel, I. G. and Clements, D. N. (2015) 'Validity of Internet-Based Longitudinal Study Data: The Elephant in the Virtual Room', *Journal of Medical Internet Research*, 17(4), p. e96. Available at: https://doi.org/10.2196/jmir.3530 (Accessed: 28 February 2022).

Quail, A., O'Reilly, C., Watson, D., McNamara, E., O'Mahony, D. and Murray, A. (2019) *A Summary Guide to Cohort '08 at 9 Years (Wave 5 of the Infant Cohort at 9 Years)*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

Quail, A., Williams, J., McCrory, C., Murray, A. and Thornton, M. (2011a) *Sample Design and Response in Wave One of the Infant Cohort (at 9 months) of Growing Up in Ireland*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

Quail, A., Williams, J., McCrory, C., Murray, A. and Thornton, M. (2011b). *A Summary Guide to Wave 1 of the Infant Cohort (9 Months) of Growing Up in Ireland*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs.

Quigley, M. A., Kelly, Y. J. and Sacker, A. (2007) 'Breastfeeding and Hospitalization for Diarrheal and Respiratory Infection in the United Kingdom Millennium Cohort Study', *Pediatrics,* 119(4), pp. e837–e842. Available at: https://doi.org/10.1542/peds.2006-2256 (Accessed: 28 February 2022).

Radin, J. M., Steinhubl, S. R., Su, A. I., Bhargava, H., Greenberg, B., Bot, B. M., Doerr, M. and Topol, E. J. (2018) 'The Healthy Pregnancy Research Program: Transforming pregnancy research through a ResearchKit app', *Npj Digital Medicine*, 1(1), p. 45. Available at: https://doi.org/10.1038/s41746-018-0052-2 (Accessed: 28 February 2022).

Raine, K., Minke, S. W., Khalema, E., Smith, C. and Plotnikoff, R. C. (2006) 'The Evolution of Integrated Chronic Disease Prevention in Alberta, Canada', *Preventing Chronic Disease*, 3(3), p. A102. Available at: https://doi:10.1177/1524839913511627 (Accessed: 28 February 2022).

Ratan, Z. A., Parrish, A. M., Zaman, S. B., Alotaibi, M. S. and Hosseinzadeh, H. (2021) 'Smartphone Addiction and Associated Health Outcomes in Adult Populations: A Systematic Review', *International Journal of Environmental Research and Public Health*, 18(22), p. 12257. Available at: https://doi.org/10.3390/ijerph182212257 (Accessed: 28 February 2022).

Raynor, P. and Born in Bradford Collaborative Group (2008) 'Born in Bradford, a cohort study of babies born in Bradford, and their parents: Protocol for the recruitment phase', *BMC Public Health*, 8(1), p. 327. Available at: https://doi.org/10.1186/1471-2458-8-327 (Accessed: 28 February 2022).

Rich, C., Cortina-Borja, M., Dezateux, C., Geraci, M., Sera, F., Calderwood, L., Joshi, H. and Griffiths, L. J. (2013a) 'Predictors of non-response in a UK-wide cohort study of children's accelerometer-determined physical activity using postal methods', *BMJ Open*, 3(3), p. e002290. Available at: https://doi.org/10.1136/bmjopen-2012-002290 (Accessed: 28 February 2022).

Rich, C., Geraci, M., Griffiths, L., Sera, F., Dezateux, C., and Cortina-Borja, M. (2013b) 'Quality Control Methods in Accelerometer Data Processing: Defining Minimum Wear Time', *PLOS ONE*, 8(6), p. e67206. Available at: https://doi.org/10.1371/journal.pone.0067206 (Accessed: 28 February 2022).

Richiardi, L., Pizzi, C. and Pearce, N. (2013) 'Commentary: Representativeness is usually not necessary and often should be avoided', *International Journal of Epidemiology*, 42(4), pp. 1018–1022. Available at: https://doi.org/10.1093/ije/dyt103 (Accessed: 28 February 2022).

Richmond, R. C., Al-Amin, A., Smith, G. D. and Relton, C. L. (2014) 'Approaches for drawing causal inferences from epidemiological birth cohorts: a review', *Early Human Development*, 90(11), pp. 769–780. Available at: https://doi.org/10.1016/j.earlhumdev.2014.08.023 (Accessed: 28 February 2022).

Ridgeway, J. L., Wang, Z., Rutten, L. J. F., van Ryn, M., Griffin, J. M., Murad, M. H., Asiedu, G. B., Egginton, J. S. and Beebe, T. J. (2017) 'Conceptualising paediatric health disparities: A metanarrative systematic review and unified conceptual framework', *BMJ Open*, 7(8), p. e015456. Available at: https://doi.org/10.1136/bmjopen-2016-015456 (Accessed: 28 February 2022).

Ries, N. M., LeGrandeur, J. and Caulfield, T. (2010) 'Handling ethical, legal and social issues in birth cohort studies involving genetic research: Responses from studies in six countries', *BMC Medical Ethics*, 11(1), p. 4. Available at: https://doi.org/10.1186/1472-6939-11-4 (Accessed: 28 February 2022).

Ritchie, F. (2017) ‘The “Five Safes”: A framework for planning, designing and evaluating data access solutions’, *Data for Policy 2017*. London, UK, 6-7 September.

Rocchi, S., Reboux, G., Frossard, V., Scherer, E., Valot, B., Laboissière, A., Zaros, C., Vacheyrou, M., Gillet, F., Roussel, S., Raherison, C. and Millon, L. (2015) ‘Microbiological characterization of 3193 French dwellings of Elfe cohort children’, *The Science of the Total Environment*, 505, pp. 1026–1035. Available at: https://doi.org/10.1016/j.scitotenv.2014.10.086 (Accessed: 28 February 2022).

Rønningen, K. S., Paltie, L., Meltzer, H. M., Nordhagen, R., Lie, K. K., Hovengen, R., Haugen, M., Nystad, W., Magnus, P. and Hoppin, J. A. (2006) ‘The Biobank of the Norwegian Mother and Child Cohort Study: A Resource for the Next 100 Years’, *European Journal of Epidemiology*, 21(8), pp. 619–625.

Rosa, E. M. and Tudge, J. (2013) ‘Urie Bronfenbrenner’s Theory of Human Development: Its Evolution From Ecology to Bioecology: The Evolution of Urie Bronfenbrenner’s Theory’, *Journal of Family Theory & Review*, 5(4), pp. 243–258. Available at: https://doi.org/10.1111/jftr.12022 (Accessed: 28 February 2022).

Sakshaug, J. W., Couper, M. P., Ofstedal, M. B. and Weir, D. R. (2012) ‘Linking Survey and Administrative Records: Mechanisms of Consent’, *Sociological Methods & Research*, 41(4), pp. 535–569. Available at: https://doi.org/10.1177/0049124112460381 (Accessed: 28 February 2022).

Salvador, P. T. C. O., Alves, K. Y. A., Rodrigues, C. C. F. M., and Oliveira, L. V. E. (2020). Online data collection strategies used in qualitative research of the health field: a scoping review. *Revista gaucha de enfermagem*, *41*, e20190297. https://doi.org/10.1590/1983-1447.2020.20190297 ///// Available at: https://doi.org/10.1590/1983-1447.2020.20190297 (Accessed: 28 February 2022).

Scherpenzeel, A., Bristle, J., Pflüger, S.-M., Butt, S., Zins, S., Emery, T., Luijkx, R., Maineri, A. and Mindarova, I. (2016) *Report on the use of sampling frames in European studies*.

Deliverable 2.1 of the SERISS Project Funded under the European Union’s Horizon 2020 Research and Innovation Programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables (Accessed: 28 February 2022).

Schmidt Morgen, C., Rokholm, B., Sjöberg Brixval, C., Schou Andersen, C., Geisler Andersen, L., Rasmussen, M., Nybo Andersen, A. M., Due, P. and Sørensen, T. I. (2013) ‘Trends in prevalence of overweight and obesity in Danish infants, children and adolescents--are we still on a plateau?’, *PLOS ONE*, 8(7), p. e69860. Available at: https://doi.org/10.1371/journal.pone.0069860 (Accessed: 28 February 2022).

Schmidt, M., Schmidt, S. A. J., Adelborg, K., Sundbøll, J., Laugesen, K., Ehrenstein, V. and Sørensen, H. T. (2019) ‘The Danish health care system and epidemiological research: From health care contacts to database records’, *Clinical Epidemiology*, 11, pp. 563–591. Available at: https://doi.org/10.2147/CLEP.S179083 (Accessed: 28 February 2022).

Schrøder, H., Rechnitzer, C., Wehner, P. S., Rosthøj, S., Møller, J. K., Lausen, B., Petersen, G. and Nørgaard, M. (2016) 'Danish Childhood Cancer Registry', *Clinical Epidemiology*, 8, pp. 461–464. Available at: https://doi.org/10.2147/CLEP.S99508 (Accessed: 28 February 2022).

Seifert, A., Christen, M., and Martin, M. (2018). Willingness of Older Adults to Share Mobile Health Data with Researchers. *GeroPsych*, 31(1), 41-49. Available at: https://econtent.hogrefe.com/doi/10.1024/1662-9647/a000181 (Accessed: 28 February 2022).

Sera, F., Griffiths, L., Dezateux, C. and Cortina-Borja, M. (2018) *Technical report on the enhancement of Millennium Cohort Study data with linked electronic health records; derivation of consent weights*. London: Centre for Longitudinal Studies, Institute of Education, University of London.

Shamsipour, M, Pirjani, R, Jeddi, MZ, Effatpanah, M, Rastkari, N, Kashani, H, Shirazi, M, Hassanvand, MS, NinoKünzli, Shariat, M, Javadi, FS, Shariatpanahi, G, Hassanpour, G, Peykarporsan, Z, Jamal, A, Ardestani, ME, Hoseini, FS, Dalili, H, Nayeri, FS, Mesdaghinia, A, Naddafi, K, Shahtaheri, SJ, Nasseri, S, Yunesian, F, Rezaeizadeh, G, Amini, H, Yokoyama, K, Vigeh, M and Yunesian, M. (2020) 'Tehran environmental and neurodevelopmental disorders (TEND) cohort study: Phase I, feasibility assessment', *Journal of Environmental Health Science & Engineering*, 18(2), pp. 733–742. Available at: https://doi.org/10.1007/s40201-020-00499-4 (Accessed: 28 February 2022).

Shepherd, P. and Gilbert, E. (2019) *Millennium Cohort Study: Ethical Review and Consent*. London: Centre for Longitudinal Studies, Institute of Education. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-Ethical-Approval-and-Consent-2019.pdf (Accessed: 28 February 2022).

Shepherd, P., Smith, K., Joshi, H. and Dex, S. (2004) *Millennium Cohort Study First Survey: A Guide to the SPSS Dataset (3rd Edition)*. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-Guide-to-the-Datasets-022014.pdf (Accessed: 28 February 2022).

Shonkoff, J. P. (2010) 'Building a New Biodevelopmental Framework to Guide the Future of Early Childhood Policy', *Child Development*, 81(1), pp. 357–367. Available at: https://doi.org/10.1111/j.1467-8624.2009.01399.x (Accessed: 28 February 2022).

Silfee, V. J., Haughton, C. F., Jake-Schoffman, D. E., Lopez-Cepero, A., May, C. N., Sreedhara, M., Rosal, M. C. and Lemon, S. C. (2018) 'Objective measurement of physical activity outcomes in lifestyle interventions among adults: A systematic review', *Preventive Medicine Reports*, 11, pp. 74–80. Available at: https://doi.org/10.1016/j.pmedr.2018.05.003 (Accessed: 28 February 2022).

Skilton, M. (2016). 'Trends of Technological Ecosystems', in *Building Digital Ecosystem Architectures. Business in the Digital Economy*. London: Palgrave Macmillan. https://doi.org/10.1057/9781137554123_2

Skopek, J. and Passaretta, G. (2021) 'Socioeconomic Inequality in Children's Achievement from Infancy to Adolescence: The Case of Germany', *Social Forces*, 100(1), pp. 86–112. Available at: https://doi.org/10.1093/sf/soaa093 (Accessed: 28 February 2022).

Smith, G., Noble, M., Goldblatt, P., Beatty, R., Bradbury, J., Dodd, K. and Shaw, A. (2007) *The Millennium Cohort Study: Technical Report on Sampling (4th Edition).* London: Centre for Longitudinal Studies, Bedford Group for Lifecourse and Statistical Studies Institute of Education, University of London.

Smith, K. and Joshi, H. (2002) 'The Millennium Cohort Study', *Population Trends*, 107, pp. 30–34.

Smyth, E. (2018) *The transition to primary education: Insights from the Growing Up in Ireland study.* Dublin: Economic and Social Research Institute. Available at: https://doi.org/10.26504/bkmnext360 (Accessed: 28 February 2022).

Smyth, E., Murray, A., McNamara, E., O'Mahony, D., Nolan, A., Duggan, B. and Shaw, J. (2022) *Growing Up in Ireland: Research Needs for Wave 5: Age 25*. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children, Equality, Disability, Integration and Youth. Available at: https://www.esri.ie/system/files/publications/BKMNEXT425_0.pdf (Accessed: 28 February 2022).

Smyth, R. L. (2009) 'Assessing the National Children's Study', *The Lancet*, 373(9661), pp. 371–372. Available at: https://doi.org/10.1016/S0140-6736(09)60120-3 (Accessed: 28 February 2022).

Spencer, N. (2010) 'Child health inequities', *Paediatrics and Child Health,* 20(4), pp. 157–162. Available at: https://doi.org/10.1016/j.paed.2009.11.002 (Accessed: 28 February 2022).

Spittle, A. J., Olsen, J., Kwong, A., Doyle, L. W., Marschik, P. B., Einspieler, C. and Cheong, J. L. Y. (2016) 'The Baby Moves prospective cohort study protocol: Using a smartphone application with the General Movements Assessment to predict neurodevelopmental outcomes at age 2 years for extremely preterm or extremely low birthweight infants', *BMJ Open*, 6(10), p. e013446. Available at: https://doi.org/10.1136/bmjopen-2016-013446 (Accessed: 28 February 2022).

Staatz, C. B., Kelly, Y., Lacey, R. E., and Hardy, R. (2021) 'Area-level and family-level socioeconomic position and body composition trajectories: Longitudinal analysis of the UK Millennium Cohort Study', *The Lancet Public Health,* 6(8), pp. e598–e607. Available at: https://doi.org/10.1016/S2468-2667(21)00134-1 (Accessed: 28 February 2022).

Stafford, M., Black, S., Shah, I., Hardy, R., Pierce, M., Richards, M., Wong, A. and Kuh, D. (2013) 'Using a birth cohort to study ageing: Representativeness and response rates in the National Survey of Health and Development', *European Journal of Ageing*, 10(2), pp. 145–157. Available at: https://doi.org/10.1007/s10433-013-0258-8 (Accessed: 28 February 2022).

Stone, C. J. and Skinner, A. L. (2017) *New technology and novel methods for capturing health-related data in longitudinal and cohort studies: Report from a CLOSER workshop*. London: CLOSER. Available at: https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-New-technology-and-novel-methods-for-capturing-health-related-data.pdf (Accessed: 28 February 2022).

Sullivan, A., Joshi, H. and Williams, J. (2020) *New birth cohort study: Theoretical sampling design options*. CLS Working Paper No. 2020/4. London: Centre for Longitudinal Studies.

Sullivan, A., Ketende, S. and Joshi, H. (2013) 'Social Class and Inequalities in Early Cognitive Scores', *Sociology*, 47(6), pp. 1187–1206. Available at: https://doi.org/10.1177/0038038512461861 (Accessed: 28 February 2022).

Sung, V., Williams, K., Perlow, E., Hu, Y. J., Ahern, S., Said, J. M., Karanatsios, B., Hopper, J. L., McNeil, J. J., Donnan, L., Goldfeld, S. and Wake, M. (2021) 'Enhancing Value and Uptake for Whole-Population Cohorts of Children and Parents: Methods to Integrate Registries into the Generation Victoria Cohort', *Children*, 8(4), p. 285. Available at: https://doi.org/10.3390/children8040285 (Accessed: 28 February 2022).

Swift, A., McConkey, R., Curry, P. and Iriarte, E. G. (2021) 'Social-Emotional Difficulties in Irish Children Aged Five and Nine Years: A National, Longitudinal Study', *Children*, 8(8), p. 656. Available at: https://doi.org/10.3390/children8080656 (Accessed: 28 February 2022).

Szymczyk, A., Pollock, G., Goswami, H., Knies, G. and Hothersall, G. (2019) *EUROCOHORT: Questionnaire Framework. Deliverable 8.3*. EuroCohort. Available at: http://www.guidecohort.eu/#:~:text=GUIDE%20(Growing%20Up%20In%20Digital,wellbeing%20of%20children%2C%20young%20people (Accessed: 28 February 2022).

Teague, S., Youssef, G. J., Macdonald, J. A., Sciberras, E., Shatte, A., Fuller-Tyszkiewicz, M., Greenwood, C., McIntosh, J., Olsson, C. A. and Hutchinson, D. (2018) 'Retention strategies in longitudinal cohort studies: A systematic review and meta-analysis', *BMC Medical Research Methodology*, 18(1), p. 151. Available at: https://doi.org/10.1186/s12874-018-0586-7 (Accessed: 28 February 2022).

Thalineau, A. and Nowik, L. (2018) 'Place des grands-parents après la naissance du nouveau-né. Les attentes des parents participant à la cohorte Elfe', *Revue des politiques sociales et familiales*, 126(1), pp. 9–20. Available at: https://doi.org/10.3406/caf.2018.3259 (Accessed: 28 February 2022).

Thompson, L., Kemp, J., Wilson, P., Pritchett, R., Minnis, H., Toms-Whittle, L., … and Gillberg, C. (2010) 'What have birth cohort studies asked about genetic, pre- and perinatal exposures and child and adolescent onset mental health outcomes? A systematic review', *European Child & Adolescent Psychiatry*, 19(1), pp. 1–15. Available at: https://doi.org/10.1007/s00787-009-0045-4 (Accessed: 28 February 2022).

Thornton, M., Williams, J., McCrory, C., Murray, A. and Quail, A. (2013) *Design, Instrumentation and Procedures for the Infant Cohort at Wave One (9 Months).* Technical Report No. 2. Dublin: Economic and Social Research Institute, Trinity College Dublin, Department of Children and Youth Affairs. Available at: https://www.growingup.ie/pubs/BKMNEXT252.pdf (Accessed: 28 February 2022).

Tienda, M. and Koffman, D. (2021) 'Using Paradata to Evaluate Youth Participation in a Digital Diary Study', *Social Science Computer Review*, 39(4), pp. 666–686. Available at: https://doi.org/10.1177/0894439320929272 (Accessed: 28 February 2022).

Tingay, K. S., Dezateux, C., Fitzsimons, E., Walton, S., Setakis, E., Cortina-Borja, M., Bedford, H., Brophy, S., Akbari, A., Griffiths, L., Bandyopadhyay, A. and Lyons, R. (2019) 'Record linkage to enhance consented cohort and routinely collected health data from a UK birth cohort', *International Journal of Population Data Science*, 4(1), Article 1. Available at: https://doi.org/10.23889/ijpds.v4i1.579 (Accessed: 28 February 2022).

Townsend, C., Cortina-Borja, M., Peckham, C., Brown, D., Joshi, H. and Dezateux, C. (2012) *Technical report on the Millennium Cohort Study biomedical data enhancement study of infections and later allergies*. London: Centre for Longitudinal Studies, Institute of Education, University of London. Available at: https://sp.ukdataservice.ac.uk/doc/6993/mrdoc/pdf/mcs2_oral_fluid_data_collection.pdf (Accessed: 28 February 2022).

Townsend, M. L., Kelly, M. A., Pickard, J. A., Larkin, T. A., Flood, V. M., Caputi, P., Wright, I. M., Jones, A. and Grenyer, B. F. S. (2019) 'Illawarra Born cross-generational health study: Feasibility of a multi-generational birth cohort study', *Pilot and Feasibility Studies*, 5(1), p. 32. Available at: https://doi.org/10.1186/s40814-019-0418-5 (Accessed: 28 February 2022).

Townsend, M. L., Riepsamen, A., Georgiou, C., Flood, V. M., Caputi, P., Wright, I. M., Davis, W. S., Jones, A., Larkin, T. A., Williamson, M. J. and Grenyer, B. F. S. (2016) 'Longitudinal Intergenerational Birth Cohort Designs: A Systematic Review of Australian and New Zealand Studies', *PLOS ONE*, 11(3), p. e0150491. Available at: https://doi.org/10.1371/journal.pone.0150491 (Accessed: 28 February 2022).

Tusla (2021) *Guidance for Research Participants Under the Age of 18.* Available at: https://www.tusla.ie/research/tusla-research-office/guidance-for-research-participants-under-the-age-of-18/ (Accessed: 28 February 2022).

UCL (2014) *Millennium Cohort Study: Building a picture of a new generation*. London: University College London. Available at: https://www.ucl.ac.uk/impact/casestudies/2014/dec/millennium-cohort-study-building-picturenew-generation (Accessed: 28 February 2022).

UCL (2016) *Why did Life Study close?* Available at: http://www.lifestudy.ac.uk/about/faqs.html (Accessed: 28 February 2022).

UCL (2019) *Research with Children: Guidance on Data Protection Issues*. London: Legal Services Data Protection Office, University College London. Available at: https://www.ucl.ac.uk/data-protection/sites/data_protection/files/research_with_children_guidance_v1.2.pdf (Accessed: 28 February 2022).

Understanding Society (2022) *Innovation Panel*. Available at: https://www.understandingsociety.ac.uk/documentation/innovation-panel (Accessed: 28 February 2022).

van Gelder, M. M. H. J., Merkus, P. J. F. M., van Drongelen, J., Swarts, J. W., van de Belt, T. H. and Roeleveld, N. (2020) 'The PRIDE Study: Evaluation of online methods of data collection', *Paediatric and Perinatal Epidemiology*, 34(5), pp. 484–494. Available at: https://doi.org/10.1111/ppe.12618 (Accessed: 28 February 2022).

Van Mol, C. (2017) 'Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response', *International Journal of Social Research Methodology*, 20(4), pp. 317–327. Available at: https://doi.org/10.1080/13645579.2016.1185255 (Accessed: 28 February 2022).

Vandentorren, S., Bois, C., Pirus, C., Sarter, H., Salines, G., Leridon, H. and the Elfe team (2009) 'Rationales, design and recruitment for the Elfe longitudinal study', *BMC Pediatrics*, 9(1), p. 58. Available at: https://doi.org/10.1186/1471-2431-9-58 (Accessed: 28 February 2022).

Vercruyssen, A., Roose, H., Carton, A. and Putte, B. V. D. (2014) 'The effect of busyness on survey participation: Being too busy or feeling too busy to cooperate?', *International Journal of Social Research Methodology*, 17(4), pp. 357–371. Available at: https://doi.org/10.1080/13645579.2013.799255 (Accessed: 28 February 2022).

Violato, M., Taylor-Robinson, D., Hungerford, D., Gray, A., O'Brien, S. and Iturriza-Gomara, M. (2019) 'Family income and exposure to norovirus in childhood: Findings from the UK Millennium Cohort Study', *SSM - Population Health*, 8, p. 100445. Available at: https://doi.org/10.1016/j.ssmph.2019.100445 (Accessed: 28 February 2022).

Wake, M., Hu, Y. J., Warren, H., Danchin, M., Fahey, M., Orsini, F., ... and Davidson, A. (2020) 'Integrating trials into a whole-population cohort of children and parents: statement of intent (trials) for the Generation Victoria (GenV) cohort', *BMC Medical Research Methodology*, 20(1), pp. 1–15. Available at: https://doi.org/10.1186/s12874-020-01111-x (Accessed: 28 February 2022).

Wallace, E., Gibb, J., Greene, V. and Joshi, P. (2013) *Millennium Cohort Study Participant Engagement Research*. London: National Children's Bureau. Available at: https://cls.ucl.ac.uk/wp-content/uploads/2018/08/MCS-Participant-Engagement-Research-National-Childrens-Bureau-2013.pdf (Accessed: 28 February 2022).

Walsh, B., Mac Domhnaill, C. and Mohan, G. (2021) *Developments in healthcare information systems in Ireland and internationally*. Economic and Social Research Institute Survey and Statistical Report Series No. 105. Dublin: Economic and Social Research Institute. Available at: https://doi.org/10.26504/sustat105 (Accessed: 28 February 2022).

Wang, J., Hu, Y. J., Clifford, S., Goldfeld, S. and Wake, M. (2021) 'Selecting life course frameworks to guide and communicate large new cohort studies: Generation Victoria (GenV) case study', *Journal of Developmental Origins of Health and Disease*, 12(6), pp. 829-848. Available at: https://doi.org/10.1017/S2040174420001245 (Accessed: 28 February 2022).

Warren, D., Quinn, B. and Daraganova, G. (2020) *Use of health services among children at risk of social-emotional problems.* Melbourne: Growing Up in Australia. Available at: https://aifs.gov.au/publications/use-health-services-among-children-risk-social-emotional-problems (Accessed: 28 February 2022).

Watson, D., Maître, B., Whelan, C. T. and Williams, J. (2014) *Growing Up in Ireland: Dynamics of Child Economic Vulnerability and Socio-Emotional Development–An Analysis of the First Two Waves of the Growing Up in Ireland Study*. Infant & Child Cohorts Research Report No. 1. Dublin: The Stationery Office.

Watson, N. and Lynn, P. (2021) 'Refreshment Sampling for Longitudinal Surveys', in P. Lynn (ed.) *Advances in Longitudinal Survey Methodology*. Essex: John Wiley & Sons, pp. 1–22.

Wenz, A., Jäckle, A. and Couper, M. P. (2019) 'Willingness to use mobile technologies for data collection in a probability household panel', *Survey Research Methods*, 13(1), p. 22.

Western, M., Haynes, M., Baffour, B. and Perales, F. (2014) *The Case for a New Australian Birth Cohort Study*. Queensland: Institute for Social Science Research, University of Queensland.

Wey, T. W., Doiron, D., Wissa, R., Fabre, G., Motoc, I., Noordzij, J. M., Ruiz, M., Timmermans, E., van Lenthe, F. J., Bobak, M., Chaix, B., Krokstad, S., Raina, P., Sund, E. R., Beenackers, M. A. and Fortier, I. (2021) 'Overview of retrospective data harmonisation in the MINDMAP project: process and results', *Journal of Epidemiology and Community Health*, 75(5), pp. 433–441. Available at: https://doi.org/10.1136/jech-2020-214259 (Accessed: 28 February 2022).

WHO (2006) *Action framework for the prevention and control of chronic diseases*. Available at: https://www.who.int/countries/moz/publications/action_framework.pdf (Accessed: 28 February 2022).

Wickman, M., Kull, I., Pershagen, G. and Nordvall, S. L. (2002) 'The BAMSE project: presentation of a prospective longitudinal birth cohort study', *Pediatric Allergy and Immunology*, 13, pp. 11–13. Available at: https://doi.org/10.1034/j.1399-3038.13.s.15.10.x (Accessed: 28 February 2022).

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3(1), pp. 1–9. Available at: https://doi.org/10.1038/sdata.2016.18 (Accessed: 28 February 2022).

Williams, J., Greene, S., McNally, S., Murray, A. and Quail, A. (2010) *Growing Up in Ireland: The infants and their families*. Infant Cohort Research Report No.1. Dublin: The Stationery Office.

Williams, J., Murray, A., McCrory, C. and McNally, S. (2013) *Growing Up in Ireland national longitudinal study of children: Development from birth to three years*. Report 5. Dublin: Department of Children and Youth Affairs. Available at: http://hdl.handle.net/10147/555696 (Accessed: 28 February 2022).

Wooldridge, J. M. (2016) *Introductory econometrics: A modern approach*. Boston, MA: Cengage Learning.

Woolfenden, S., Eapen, V., Axelsson, E., Hendry, A., Jalaludin, B., Dissanayake, C., Overs, B., Descallar, J., Eastwood, J., Einfeld, S., Silove, N., Short, K., Beasley, D., Črnčec, R., Murphy, E. and Williams, K. (2016) 'Who is our cohort: Recruitment, representativeness, baseline risk and retention in the "Watch Me Grow" study?', *BMC Pediatrics*, 16(1), p. 46. Available at: https://doi.org/10.1186/s12887-016-0582-1 (Accessed: 28 February 2022).

Yang, K., Petersen, K. J. and Qualter, P. (2020) 'Undesirable social relations as risk factors for loneliness among 14-year-olds in the UK: Findings from the Millennium Cohort Study', *International Journal of Behavioral Development*, 46(1), 3–9. Available at: https://doi.org/10.1177/0165025420965737 (Accessed: 28 February 2022).

Yang, X., Jiao, R., Yang, L., Wu, L.-P., Li, Y.-R. and Wang, J. (2011) 'New-generation high-throughput technologies based "omics" research strategy in human disease', *Yi Chuan = Hereditas*, 33(8), pp. 829–846.

Zhu, J. L., Hjollund, N. H., Andersen, A.-M. N. and Olsen, J. (2004) 'Shift work, Job Stress, and Late Fetal Loss: The National Birth Cohort in Denmark', *Journal of Occupational and Environmental Medicine*, 46(11), pp. 1144–1149. Available at: https://doi.org/10.1097/01.jom.0000145168.21614.21 (Accessed: 28 February 2022).

Zhu, J. L., Olsen, J., Liew, Z., Li, J., Niclasen, J. and Obel, C. (2014) 'Parental Smoking During Pregnancy and ADHD in Children: The Danish National Birth Cohort', *Pediatrics*, 134(2), pp. e382–e388. Available at: https://doi.org/10.1542/peds.2014-0213 (Accessed: 28 February 2022).

Zhu, Y., Olsen, S. F., Mendola, P., Halldorsson, T. I., Rawal, S., Hinkle, S. N., Yeung, E. H., Chavarro, J. E., Grunnet, L. G., Granström, C., Bjerregaard, A. A., Hu, F. B. and Zhang, C. (2017) 'Maternal consumption of artificially sweetened beverages during pregnancy, and offspring growth through 7 years of age: a prospective cohort study', *International Journal of Epidemiology*, 46(5), pp. 1499–1508. Available at: https://doi.org/10.1093/ije/dyx095 (Accessed: 28 February 2022).

Zylbersztejn, A., Gilbert, R. and Hardelid, P. (2020) 'Developing a national birth cohort for child health research using a hospital admissions database in England: The impact of changes to data collection practices', *PLOS ONE*, 15(12), pp. 1-17. Available at: https://doi.org/10.1371/journal.pone.0243843 (Accessed: 28 February 2022).

www.growingup.gov.ie