

The SA Early Childhood Data Project

As part of the *BetterStart* Research Series this report provides an overview of the Early Childhood Data Project. It includes descriptions of the various datasets and highlights the opportunity the linked data presents for informing optimal child health and development.

The *BetterStart* Child Health and Development Research Group comprises interdisciplinary researchers from epidemiology, public health, nutrition, paediatrics, biostatistics, and psychology who are trying to better understand how to ensure infants and children have the best start in life that will enhance their health, development and human capability formation over the life course.

For further information contact
Professor John Lynch
john.lynch@adelaide.edu.au
<https://health.adelaide.edu.au/betterstart/research/>

The South Australian Early Childhood Data Project (ECDP) is one of the most comprehensive population-based administrative research databases in Australia. It spans more than 30 different government administrative data sources and has been built over a number of years to include every birth cohort of South Australian children born from 1999-2013. The value of the ECDP is that it is a resource that can inform research, service provision and policy around child health, development and human capability formation from the perinatal period into adolescence.

The ECDP has been a platform for a number of research and academic partnerships and this has included work with the South Australian Department of the Premier and Cabinet, Child and Family Health Service, SA Health, Department for Education and Child Development, Families SA, Women's and Children's Health Network, Department for State Development, The Council for the Care of Children, Wardliparingga Aboriginal Research Unit, and the Aboriginal Health Council.

Currently, 37 approved researchers from universities and government departments in South Australia have access to the ECDP research platform. Together we work to build and use a public-good data resource that supports the best start in life for all South Australian children and enhances their health and development throughout their lifecourse.

Introduction

Childhood is a time of rapid development and it is critical to understand and prevent exposures that adversely affect the health and wellbeing of children from the perinatal period to adolescence¹. Understanding how children can fulfil their health and development potential is important for their ability to benefit from learning opportunities, academic achievement at school and for their future health, social and economic capability^{2,3}. The key objective of the ECDP is to link administrative data sources for every cohort of children in South Australia (SA) who were born from 1999-2013, to

explore the effects that early life conditions and experiences have on child development, health and learning. The ECDP has developed through extensive human research ethics and custodian approvals, data acquisition processes and academic partnerships to deliver a public-good linked data platform which aims to benefit all children, families and communities in South Australia.



Project background

Acquisition of physical, social, emotional and cognitive capacities between birth and age eight occurs at a rate which exceeds that during any other stage of life.^{4,5} During this stage of life, children undergo many developmental milestones and key social transitions such as entry into childcare, preschool and school. Recognition that human wellbeing, health, cognitive skills and social competence underpins economic development⁶ has helped drive the argument for investment in early childhood, as trajectories of later life outcomes are strongly influenced by early experiences.⁷ Ensuring that all children have the best start in life is a priority in national⁸ and state⁹ policies.

The use of linked de-identified administrative data provides opportunities for examining a broad range of child health and development outcomes that align with those selected in national priorities as key performance indicators. The research proposed for the ECDP allows a whole-of-population lens to focus on the whole child, across physical health, social, emotional, cultural, cognitive and learning dimensions from before birth to adolescence.

The aims of the ECDP currently include:

- Investigate the consequences of pregnancy complications, and poor outcomes at birth on child health, vulnerability at school entry, and school achievement;
- Investigate the patterning and consequences of social, economic and demographic disadvantage on vulnerability at school entry, and transitions to school achievement;
- Describe the epidemiology of hospital emergency presentations and potentially preventable hospitalisations (PPHs) for children. Improve understanding about the relationship between key perinatal and sociodemographic factors, and hospital emergency presentations and PPHs;
- Develop data-driven risk prediction models to inform more accurate identification of families whose children are at high risk of poor health and development, and who may benefit from intensive family support programs;

- Assess the long term functional outcomes of traumatic brain injury in relation to vulnerability at school entry and school achievement;
- Examine the associations between dental health and vulnerability at school entry, and school achievement;
- Investigate the social factors related to children in contact with the child protection system;
- Determine if children vulnerable to maltreatment can be identified early to enable targeted prevention programs;
- Investigate the developmental, academic and behavioural outcomes of maltreated children to identify any protective factors that will improve outcomes for them;
- Utilise knowledge gained from the aims above to obtain a better understanding of the social and health origins of poor academic achievement in school.

The core goal of this research is to inform better service delivery and policy across the government and non-government sectors that support the health and development of children.

Funding support

We are grateful for the funding support from:

- National Health and Medical Research Council (NHMRC) Australia Fellowship (570120);
- NHMRC Partnership Project Grant (APP1056888);
- NHMRC Centre for Research Excellence (APP1099422)

As well as contributions from:

- SA Department of the Premier and Cabinet;
- SA Department of Health;
- SA Department for Education and Child Development; and
- Australian Research Alliance for Children & Youth (ARACY).

Acknowledgments

We would like to thank SA NT DataLink and all of the data custodians and data managers from all our government partners at State and Federal levels. The views expressed here do not necessarily reflect those of our government partners.

Partners



Government of South Australia

Department of the Premier and Cabinet



Government of South Australia

Department for Education and Child Development



Government of South Australia

Department for Communities and Social Inclusion



Government of South Australia

Attorney-General's Department



Government of South Australia

Consumer and Business Services



Government of South Australia

SA Health



Australian Government

Department of Education

WOMEN'S AND CHILDREN'S HEALTH NETWORK

Child and Youth Health



Aboriginal Health Council

of South Australia Inc.

'Our health, our choice, our way'



WARDLIPARINGGA

SA·NT



DataLink

Ethics

Gaining initial ethics approval for both the data linkage and research on the linked data was an extensive process that took more than two years. It is now a continuing process of engagement with various ethics committees and research governance officers around the state, regularly updating them on progress and research objectives. Ethics approval has been granted from the South Australian Department of Health Human Research Ethics Committee (HREC) (377/06/2013; HREC/13/SAH/106), the University of Adelaide HREC (H-185-2011), the Aboriginal Health Research Ethics Committee (04-11-405; 04-13-538), the Women's and Children's Hospital Research Ethics Committee (REC2411/9/14) and the Flinders University Social and Behavioural Research Ethics Committee (5504).

The linkage process

SA NT DataLink is an unincorporated collaborative joint venture involving partners from all South Australian Universities, the South Australian Health and Medical Research Institute, the SA Department of Health, the Department for Education and Child Development, the Department for Communities and Social Inclusion, the Cancer Council SA-Beat Cancer Project and the Northern Territory Government.

The role of SA NT DataLink is to support research, policy development and planning by providing de-identified data linkage services. Its privacy protection policies have been endorsed by the SA Privacy Commission.

After progressing through an application process with SA NT DataLink, reaching agreement with data providers and custodians and obtaining ethics approval, data custodians provide the requested data to SA NT DataLink with a local record ID and key identifying variables. This information is used with the Master Linkage File by SA NT DataLink to probabilistically match the identifying information available in multiple datasets. Once matched, a project specific ID is generated and data custodians use this and their local record ID to link back to their detailed data, remove identifying information and deliver the de-identified data to the researchers with requested variables and the Project Specific Linkage Key that is really just a random ID specific to that research project. For a more

detailed explanation of the linkage process please see <https://www.santdatalink.org.au/> or view an explanatory video from SA NT DataLink found here: <https://www.youtube.com/watch?v=vLYGcbxriPA>

As unique identification numbers (i.e. for health care cards) are not used in Australia, linkages are necessarily probabilistic and are based on key demographic information and therefore a small degree of error is to be expected. Linkage error can occur from missed links or incorrect positive links.¹⁰ Although calculation of false linkages has not been undertaken in SA to date, Western Australia and New South Wales use similar systems and estimate false positive linkage error of approximately 0.1%¹¹ and 0.3%,¹² respectively.

Data validation

The ECDP team has engaged in extensive data validation and checking, involving close collaboration with custodians. Once we receive the de-identified data from the data custodian, a process of validation is commenced. Initial validation checks involve exploration of individual datasets to check that:

- data items correspond to the original approved requests;
- there is no unexplained missing or additional data;
- no duplicates exist within and between datasets;
- project-specific linkage keys match across different datasets; and
- prevalence estimates from key data items are comparable to existing publicly available reports (i.e. SA Health Pregnancy Outcomes reports, ABS reports). In cases where there are no publicly available reports, published research manuscripts utilising the same data have been used as the comparison reference.

The data validation process is ongoing and will continue to involve consultation with data custodians to ensure the highest level of validity is achieved before any research is conducted using the information.

Data access

The ECDP has been ethically approved and built as a public-good resource to assist Government and research collaborators to use the data in new and better ways. At the same time, the ECDP aims to

address questions of relevance to advancing the science of early childhood health and development. All researchers with access to the linked data have entered into deeds of confidentiality with various organisations and been approved by the SA Health HREC.

Approved external researchers access the ECDP via a secure online platform provided by the University of Adelaide.

Data description

The data currently linked for the ECDP includes multiple datasets from the Births and Deaths Registry, Perinatal statistics, the Birth Defects Registry, Public Hospital Inpatient and Emergency Presentation data, Congenital Abnormality data, Child Health Records, Family Home Visiting Program, the Edinburgh Postnatal Depression Scale (EPDS), Postnatal Risk Questionnaire (PNRQ), the Brief Response, Pathways to Parenting Questionnaire, Childhood Immunisations, Neonatal Hearing, Public School Enrolment Census, Reading achievement levels, English as an Additional Language/Dialect (EAL/D), the National Assessment Program Literacy and Numeracy (NAPLAN), the Australian Early Development Census (AEDC), Child Protection, Youth Justice, Public Housing and the SA Titanium Dental data. We are also in negotiations with multiple custodians to receive additional data into the future to enable a richer characterisation of child health and development in South Australia.

Utilising variables from SA Births Registry data, a 'Family File' has been created which links siblings with a registered birth and born to the same mother between 1999-2013. This file enables researchers to analyse the social, emotional and physical health relationships that occur within families. The 'backbone' of all data analysis conducted by ECDP researchers draws on the 'Born Population File'. This file is a merge of the Births and Perinatal datasets which represents our best estimate of population denominators for all birth cohorts for SA children born 1999-2013.

Figure 1 ECDP Data Themes

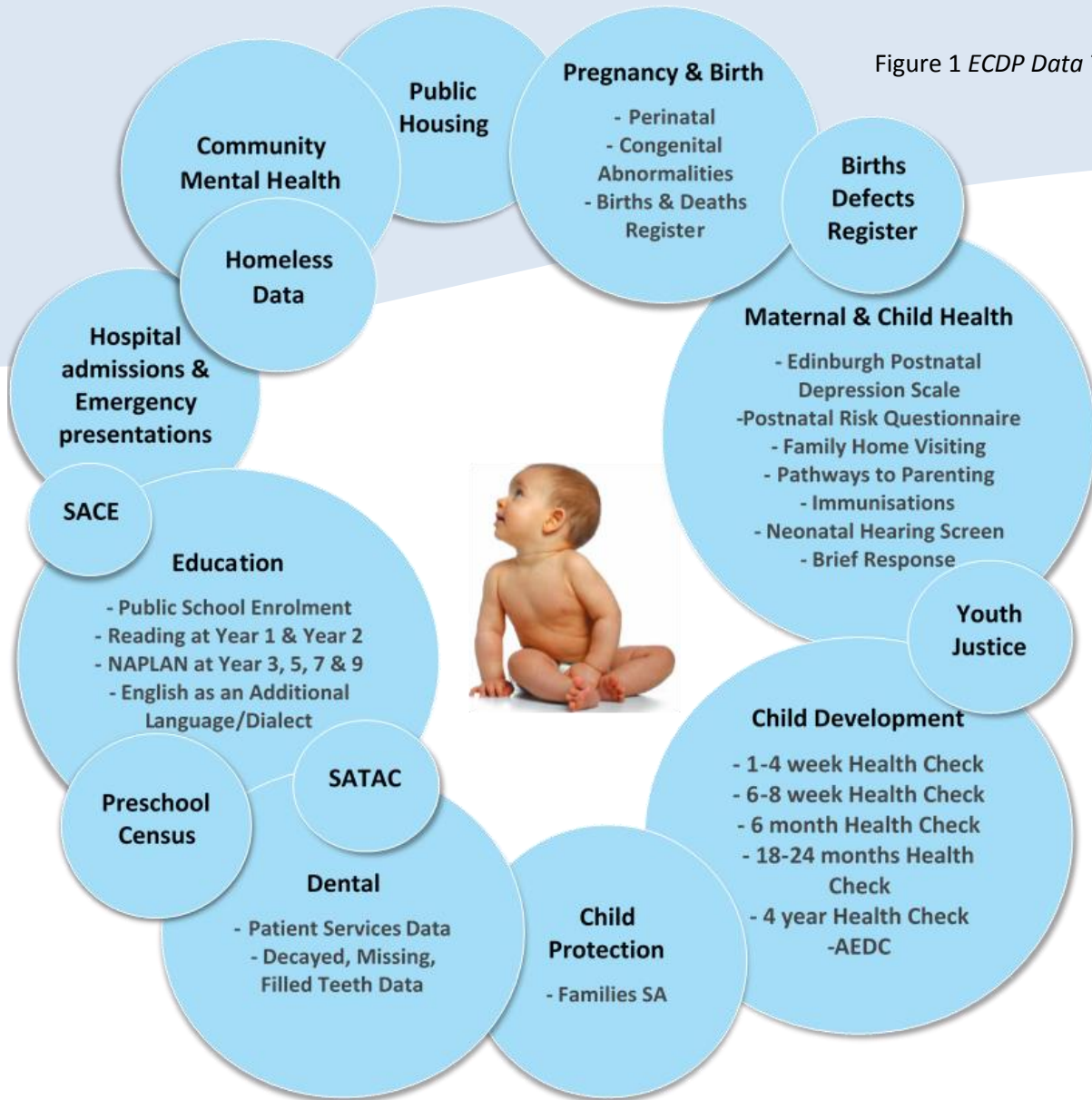


Figure 1 above illustrates the themes and datasets currently part of the ECDP, and those that are under negotiation to be added to the project (SACE, SATAC, Homeless, and Community Mental Health). Table 1 (page 14) details which datasets are currently available for specific birth years and age ranges.

Please note that all figures provided in this report are rounded estimates of the n available in a particular dataset. Once any two datasets are merged, these n 's obviously change. The estimates provided here represent the number of individuals within a specific dataset, as opposed to number of records. For example, one individual may have 10 records of hospitalisations in the ISAAC dataset but will only be counted as one individual. Indeed, this is one of the key advantages of linked data - it turns episodes of care into individual experiences. Specific

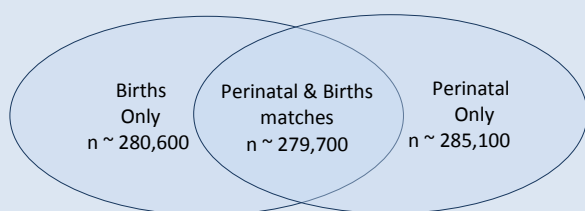
sample sizes will vary depending on datasets and variables linked for specific birth cohort years. In addition, although the terms parental and maternal are used throughout this report, this may also refer to primary caregiver(s) where appropriate.

BORN POPULATION FILE

A synthetic dataset compiled by the ECDP Group

The ECDP team has merged the Births and Perinatal datasets to create what is known as the 'Born Population'. As it is possible that some births are missed in the Births registration or Perinatal data collection process, matching these two datasets has created the most accurate representation of all children born in South Australia between 1999 and 2013. Figure 2 depicts the matches and mismatches between the two datasets.

Figure 2. Perinatal dataset and Births dataset merged



This dataset serves as the core dataset for linkage to every other dataset that is a part of the ECDP, as a significant amount of work has been undertaken to verify and establish this file as the most reliable source of demographic characteristics. In addition to this, as data is collected by different agencies it has been necessary to develop standard formats for common data items that may originally have been coded differently (i.e. gender format).

As with any data linkage process, there are potential sources of error that may lead to inconsistencies for similar variables between datasets such as reporting, recording, data-entry, data-management error and missed or false positive linkages. Extensive checks have been undertaken to create a set of what is referred to as 'common variables' that have been verified and deemed reliable based on information from multiple datasets. The core common variables include mother and child date of birth, child sex, infant plurality, birth order, birth weight and postcode at time of birth and a variable that used multiple datasets to determine the Aboriginal ethnicity of the child¹³. The ECDP team have invested significant time into creating these variables and they are our 'best estimate' of the correct data. In every case we have explicit rules for creating these variables that are available to any researcher so they can judge whether they want to use our recommended variables or create their own version.

This data checking and cleaning process has involved assessing consistency of variables between datasets and developing decision rules to establish these common variables. For example, consultation with the Aboriginal Health Council of SA has resulted in a set of decision rules that uses information from both the Birth and Perinatal data for people of Aboriginal or Torres Strait Islander origin¹³. Inconsistencies for other variables such as sex or gender have also been resolved through the development of decision rules based on

best linkage practice¹³. Therefore, the Born Population file contains the most accurate and reliable estimate of the core common variables.

Data

The ECDP group recommend that the core characteristics developed in the Born Population File serve as the denominator for linkage to all other datasets, as the validity and reliability of these characteristics has been verified as well as we are able to do that. However, the Born Population file contains information only on children born in SA, other datasets also contain information on children not included in this core dataset. For instance, there could be a hospitalisation for a child who is in SA on holiday from another jurisdiction or who was not born in SA.

PREGNANCY AND BIRTH THEME:

Data relating to the Pregnancy and Birth theme includes Perinatal statistics, Congenital Abnormalities and Births Registry data. The range of factors available in this theme allow us to examine the associations between antenatal health and poor outcomes at birth on child health and development, and vulnerability at school entry. The following provides detail about the datasets that we have received relating to this theme.



BIRTHS DATA

The Births data is collected as a part of the Births, Deaths and Marriages registry. All births (alive or stillborn) are legally required to be registered within 60 days of birth. Deaths that occur after 20 weeks gestation or >400 grams are also recorded in the Births Register. Available variables from this dataset include mother, father and child demographics and basic clinical birth data.

Registered Births data is currently available for every cohort of SA children born from 1999 to 2013. Approximately ~n =280,600 registered births were recorded in SA for the birth cohorts born from 1999 to 2013.

SA PERINATAL STATISTICS COLLECTION

The South Australian Perinatal Statistics Collection is a collection of information sourced from the Supplementary Birth Record completed by midwives and neonatal nurses after hospital and home births. Data is collected on all births in SA, including from women who normally live interstate. Specific information on children includes gestation at birth, child sex, birth weight and APGAR scores at one and five minutes. Further information includes maternal age; gestational health; parity; marital status; ethnicity and smoking status. Parental occupation at time of birth is also collected.

The range of factors included in the perinatal data allows examination of some maternal factors and perinatal exposures as predictors of child health and development.

Data

South Australian perinatal data is currently available for children born from 1999 to 2013. Approximately ~n =285,100 individual births are represented in the Perinatal Statistics collection from 1999 to 2013.

CONGENITAL ABNORMALITIES/ BIRTH DEFECTS REGISTER

Information on congenital abnormalities detected at birth or in the neonatal period (within 28 days of birth) is provided by practitioners using the Congenital Abnormality Form. Birth defects that are detected and notified after discharge from hospital and up to five years post birth are reported annually by the South Australian Birth Defects Register at the Women's and Children's Hospital.

Data

An indicator for the presence of specific birth defects/ congenital abnormalities in the child, the age at diagnosis and a record of any family history of congenital abnormalities is included in these datasets. This data is currently available for SA birth cohorts born 1999-2011.

MATERNAL AND CHILD HEALTH THEME:

The data included in this theme provides information on important health, social and maternal factors that may affect child health and development. The ECDP has data that includes the Pathways to Parenting Questionnaire, the Family Home Visiting Program, the Edinburgh Postnatal Depression Scale, the Postnatal Risk Questionnaire, Childhood Immunisations and the Brief Response.

PATHWAYS TO PARENTING QUESTIONNAIRE

The Pathways to Parenting questionnaire was a tool used by Child and Family Health Services (CaFHS) nurses undertaking the Universal Contact Visit following the birth of a baby. It was based on a supportive relationship between the nurse and parent(s) of the infant to assist families in identifying any needs they may have that could impact on the creation of a secure attachment with their child. Information contained in the Pathways to Parenting data includes: type of antenatal care received, mothers education and perceived access to social support.

Data

Data from the Pathways to Parenting questionnaire is available from 2005 to 2008, corresponding to the cohorts of SA children born in these years. Approximately ~n = 48,000 individuals have information in the Pathways to parenting dataset.



SOUTH AUSTRALIAN FAMILY HOME VISITING PROGRAM

The South Australian Family Home Visiting Program (FHV) is a 2-year postnatal home visiting program delivered by nurses to families requiring additional support. After a pilot phase in 2004-2005, FHV began in metropolitan SA in 2008 and was progressively rolled out to rural areas in 2010¹⁴. The FHV program aims to improve the quality of mother-child relationships; provide anticipatory guidance about infant health, safety and development; and better connect families to local community supports.

This data provides insight into some of the social, emotional and cultural factors that may impact child health and development.

Data

Data from the Family Home Visiting Program is available from 2004 until 2013. Approximately ~n =10,000 individuals exist in the Family Home Visiting program data for this period, corresponding to children born in 2004 to 2013.

THE NATIONAL PERINATAL DEPRESSION INITIATIVE

The National Perinatal Depression Initiative (NPDI) is a universal screen for depression and anxiety for women during the antenatal and postpartum periods. It improves the prevention and early detection of antenatal and postnatal depression and provides an indicator that support and treatment may be needed for expectant and new mothers experiencing depression and/or anxiety. Data relating to the NPDI draws from two questionnaires: The Postnatal Risk Questionnaire (PNRQ) and the Edinburgh Postnatal Depression Scale (EPDS).

POSTNATAL RISK QUESTIONNAIRE

The Postnatal Risk Questionnaire is a perinatal screening tool for anxiety and depression that is completed by a clinician. It asks the expectant mother thirteen questions, which are rated on a scale of 1-5 to determine maternal emotional health and any need for appropriate support services.

EDINBURGH POSTNATAL DEPRESSION SCALE

The Edinburgh Postnatal Depression Scale is a 10 item self-report measure that is designed to screen women for current symptoms of depression during pregnancy and the postnatal period. It does not diagnose depression, but is used as a tool to raise awareness of mood problems that may need further exploration or follow up.

Data

Federal funding for the NPDI was from 2008-2013, although data is still being collected in SA. Currently available data corresponds to mothers of children born within the federal funding period, where approximately ~n =13,000 individuals are represented.

CHILDHOOD IMMUNISATION

Childhood vaccinations at Child and Family Health Services (CaFHS) in SA are recorded at birth, 6 weeks, 4 months, 6 months, 12 months, 18 months, 3.5-4 years and 13 years. Immunisation data collected includes vaccine type, the date the vaccine was administered, as well as a record of any adverse reactions the child may have to the vaccine.

Data

Although we acknowledge that this data does not reflect population coverage of immunisations in SA, it does provide information about how South Australian children are immunised. Approximately ~n =282,000 individuals have at least one childhood immunisation recorded by CaFHS.

NEONATAL HEARING SCREENING PROGRAM

The Neonatal Hearing Screening program is a universal statewide service that is coordinated by the Women's and Children's Health Network in SA. The program provides free hearing screening to all newborn babies to ensure the early detection and implementation of intervention strategies for hearing loss. Babies are screened by a midwife or a designated screener before leaving hospital. If a pass result is not obtained in one or both ears, a second screen will be offered at CaFHS, with referral to an audiologist if required within the first few weeks following birth.

Data

Data available from the Neonatal Hearing Screening Program includes the Universal Hospital Screening and the CaFHS First and Second Hearing Screening. We have Universal Hospital Hearing Screening data for approximately $n = 154,000$ individuals for the birth cohorts born 2005-2013.

BRIEF RESPONSE

The Brief Response includes data recorded by health practitioners in conjunction with parents regarding child health and development issues of concern. It includes descriptive variables for the health or development issue of concern, which may relate to: breast feeding, settling, growth or immunisation, as well as family and social factors relating to child well-being. This dataset also includes any referrals to further supports that have been recommended.

Data

Brief Response data identifies potential risk factors that may affect child health and development. Approximately $n = 29,000$ individuals are captured in this dataset covering SA birth cohorts born 2005-2014.

CHILD DEVELOPMENT THEME:

This theme covers all child development data that are recorded as part of the Child Health Record. Although recorded when a child is in the school system, the AEDC data also falls under this theme as it provides information regarding a child's cognitive and social-emotional development at school entry. Data from the AEDC and Child Health Record provide valuable information about healthy child development.

CHILD HEALTH RECORD

The Child Health Record is commonly known in SA as the 'Blue Book'. This resource, developed by the Women's and Children's Health Network, is given to parents to record developmental milestones and other important information at birth and various ages up to 5 years of age for segments of the population accessing CaFHS services. Information recorded in the book can include: child demographics, growth trajectories, feeding, oral health, medical and health issues as well as the health check date. Additionally, Body Mass Index,

visual acuity, audiometry and development level are available for the preschool health check (approximately 4 years of age). The Child Health Record also includes information from the 1 to 4 week health check, undertaken at the Universal contact visit (UCV). The UCV is a service offered to all families in SA following the birth of a baby. It enables any family, child development and health issues to be identified early so that optimal child development may be promoted through early access to child health services, parenting information and support pathways for families.

Data

Child Health Record data is currently available from 1999 to 2013 corresponding to the birth cohorts of SA children born 1999 to 2013. Approximately $n = 219,000$ individuals have Child Health Record data for the 1-4 week health check; $n = 35,000$ individuals are represented in the 6-8 week health check; $n = 102,500$ individuals have data for the 6 month health check; approximately $n = 69,500$ individuals have data in the 18 month health check; $n = 39,000$ individuals are represented in the 2-3 year old health check and $n = 156,000$ individuals have data in the preschool health check. We also have approximately $n = 148,000$ individuals represented in the Universal Contact Visit data.

AUSTRALIAN EARLY DEVELOPMENT CENSUS

The Australian Early Development Census (AEDC) is funded by the federal government and is carried out across Australia once every three years (www.aedc.gov.au). This national data collection was first implemented in 2009 and involved teachers completing the AEDC questionnaire for over $n = 261,000$ students across Australia in their first year of school. The AEDC includes 95 questions that assess the following domains: physical health and wellbeing; social competence; emotional maturity; language and cognitive skills; communication skills and general knowledge. Data items available include child demographics, class type, terms attended, special needs, language, school absences, Socio-Economic Indexes for Areas (SEIFA), and individual question responses as well as domain, sub-domain scores and indicators of vulnerability. Analyses focus on factors

associated with vulnerability and how AEDC predicts future child health and school achievement.

Data

AEDC data is currently available for 2009 and 2012 and we are due to receive 2015 data shortly. These data correspond respectively with the cohorts of SA children born from 2003 to 2004, 2006 to 2007 and 2009 to 2010. Approximately $n = 42,000$ individuals have been captured in the 2009 and 2012 AEDC data.

CHILD PROTECTION THEME:

Data related to the Child Protection theme allows us to investigate a number of the social factors related to children that come in contact with the child protection system. This data provides valuable information that may help determine if children vulnerable to maltreatment can be identified early. By linking this data with other datasets, we are also able to explore the health, developmental, academic and behavioural outcomes for these children.

Data

The information in these datasets relates to de-identified reports of child abuse and neglect in South Australia. This includes children who are the subject of child protection notifications, placement in out-of-home care and/or placement under the guardianship or custody of the Minister for Families and Communities.

EDUCATION THEME:

The Education theme includes data that is available for children while they are a part of the education system. Variables include those from the Public School Enrolment Census, the English as an Additional Language/Dialect tool, Reading Achievement Records, and NAPLAN. This data allows research into language and reading achievement trajectories, cognitive skills, physical health, social and emotional wellbeing and general academic competence.

SOUTH AUSTRALIAN SCHOOL ENROLMENT CENSUS

The School Enrolment Census collects data on students enrolled and attending government schools in SA. A census is undertaken in Term 1 and Term 3 of each year and creates a snapshot of enrolment information to

assist in reporting and monitoring. Information collected includes child and parental demographics such as: parental education level and occupation, main language spoken at home, and country of birth.

Additionally, information regarding reasons for leaving school (if relevant), school absences and behaviour management incidents are recorded.

The School Enrolment Census information allows investigation of child and parental factors that may serve as predictors of future health and development.

Data

School Enrolment data is currently available from 2005 to 2013, corresponding to the birth cohorts of SA children born from 1999 to 2008. Approximately $n = 194,000$ individuals exist in the Public School Enrolment data.

ENGLISH AS AN ADDITIONAL LANGUAGE / DIALECT (EAL/D)

The Australian Curriculum, Assessment and Reporting Authority have developed a teacher resource to record EAL/D learning progression. This tool supports targeted teaching, assessment and reporting of EAL/D students. Dependent on performance across written, oral and multimodal English pieces, a child is assigned a Language and Literacy level. Previously known as the ESL scales, this level corresponds to a standard level of English required to achieve South Australian Curriculum Standards at different year levels. The Language and Literacy Levels describe the level of command an individual has of the English language, and highlights students who need targeted support i.e. if they are operating below their year level. Data items available include child demographic information, the date the Language and Literacy Levels were completed and the Language and Literacy Levels achieved for students with EAL/D.



The EAL/D data describes the development of Standard Australian English required to meet the demands of the Australian Curriculum across the years of schooling from Foundation (Reception) to Year 10. It can be used to inform targeted programming and planning to support school achievement and child development

Data

EAL/D data is currently available from 2005 to 2015 for the cohorts of SA children born from 1999 to 2009. A total of ~n= 42,000 individuals are represented in the EAL/D dataset.

READING ACHIEVEMENT (Running Records)

Reading Achievement records are a tool used by teachers to observe a student's strengths, weaknesses and development in reading proficiency through the oral assessment of a student reading of a continuous body of text. The assessment is delivered by teachers to students in Year 1 and 2. Statewide data is collected twice a year from one optional and one compulsory assessment period. Data available includes date of assessment, text level and accuracy rate.

Running Records provide information about child development and achievement in reading which can be conceptualised as an outcome or a predictor of later school achievement.

Data

Reading Achievement records data are currently available from 2008 to 2015 corresponding to the birth cohorts of SA children born from 2002 to 2009 for Year 1 tests and from 2001 to 2008 for Year 2 tests. Approximately ~n =105,000 individuals are represented in the Reading Achievement dataset for the birth cohorts born from 2001 to 2009.

NATIONAL ASSESSMENT PROGRAM-LITERACY AND NUMERACY (NAPLAN)

The NAPLAN is a national assessment of Australian students in literacy and numeracy during Years 3, 5, 7 and 9 at school, undertaken annually since 2008. Child academic achievement is assessed across four domains of reading, writing, language (grammar and punctuation) and numeracy. Specific information in this dataset is currently available for tests undertaken by

students in Year 3, 5, 7 and 9, and includes student demographics, and the index of educational disadvantage. General test results by domain as well as achievement against proficiency bands and the national minimum standard are also included.

NAPLAN data provides key indicators of school achievement.

Data

Overall, the ECDP have data relating to ~n =119,000 individuals with NAPLAN data. NAPLAN records for Year 3 assessments are currently available for 2008 to 2015 and correspond to the birth cohort years of 2000 to 2007. Approximately ~n =101,000 individuals are represented in this data. Records for Year 5 assessments are available for 2009 to 2015 and correspond with the birth cohort years of 1999 to 2005. There are approximately ~n =78,000 individuals in the year 5 NAPLAN data. Records for Year 7 assessments are currently available for 2011 to 2015 and correspond with the birth cohort years of 1999 to 2003. Data is available for approximately ~n =51,000 individuals sitting NAPLAN in year 7. Records for year 9 assessments are available for 2012 to 2015 and correspond with birth cohort years 1999 to 2001, where approximately ~n =26,000 individuals have data available.

HOSPITAL THEME:

Data found under this theme is available from SA Emergency Department presentations and episodes of Inpatient Hospital stays, known as 'separations'. By linking Hospital data we are able to investigate patterns and trends in Potentially Preventable Hospital admissions for children to inform more effective primary care and prevention.

SOUTH AUSTRALIAN PUBLIC HOSPITAL INPATIENT DATA (ISAAC)

The Integrated South Australian Activity Collection (ISAAC) is a data collection system recording information on all patients admitted to public hospitals in SA since the 1st of July, 2001. Administrative data are routinely recorded by hospital staff and updated at the time of separation i.e. discharge, transfer or death. Specific information is available for child demographics,

admission time and category, diagnoses (ICD codes) and other clinical indicators, length of stay and nature of separation.

Data

ISAAC records are currently available from the 1st of July 2001 to the 30th of June, 2014 which is inclusive of records for the birth cohorts born from 2001 to 2014. Approximately $n = 159,000$ individuals are represented in the ISAAC data for this period.

SOUTH AUSTRALIAN EMERGENCY DEPARTMENT DATA COLLECTION (EDDC)

The EDDC details emergency department presentations for all major metropolitan public hospitals in SA. Data is available from 1st July, 2003 and includes only physical presentations. Telehealth consultations are not in scope. The aim of the data collection is to demonstrate performance and patterns in Emergency presentations to aid service planning and funding. Data items include child demographics, presentation and departure information, and clinical indicators.

Data

Currently available EDDC data corresponds to the birth cohorts born from 2003 to 2014. Approximately $n = 232,000$ individuals are represented in the EDDC data.

BetterStart Publications

1. Pilkington R, Scalzi D, Gialamas A, Smithers L, Chittleborough C, Lynch J. The South Australian Early Childhood Health and Development project. Research Series 2014-01. Adelaide: School of Public Health, The University of Adelaide. 2014.
2. Sawyer A, Gialamas A, Sawyer M, Lynch J. Five by Five: A Supporting Systems Framework for Child health and Development. BetterStart Child Health and Development Research Group, School of Public Health, The University of Adelaide. 2014
3. Chittleborough C, Lynch J. Can linked data help to better target family support programs for child health and development? BetterStart Child health and Development Research Group, School of Public Health, The University of Adelaide. 2014
4. Smithers L, Gialamas A, Scheil W, Brinkman S, Lynch J. Anaemia of pregnancy, Perinatal Outcomes and Children's Developmental Vulnerability: A whole-of-population study. *Paediatric and Perinatal Epidemiology*. 2014; 28 (5): 381-390. DOI: 10.1111/ppe.12149
5. Smithers L, Searle A, Chittleborough C, Scheil W, Brinkman S, Lynch J. A Whole-of-population study of term and post term gestational age at birth and children's development. *International Journal of Obstetrics and Gynaecology*. 2015; 122 (10): 1303-11. DOI: 10.1111/1471-0528.13324
6. Chittleborough C, Searle A, Smithers L, Brinkman S, Lynch J. How well can poor child development be predicted from early life characteristics?: A whole-of-population data linkage study. *Early Childhood Research Quarterly*. 2016; 35: 19-30. DOI: 10.1016/j.ecresq.2015.10.006
7. Gialamas A, Pilkington R, Berry J, Scalzi D, Gibson O, Brown A, Lynch, J. Identification of Aboriginal children using linked administrative data: Consequences for measuring inequalities. *Journal of Paediatrics and Child Health*. 2016; 52 (5): 534-540. DOI: 10.1111/jpc.13132.
8. Pearce A, Scalzi D, Lynch J, Smithers L. Do thin, overweight and obese children have poorer development than their healthy-weight peers at the start of school? Findings from a South Australian data linkage study. *Early Childhood Research Quarterly*. 2016; 35: 85-94. DOI: 10.1016/j.ecresq.2015.10.007.
9. Smithers L, Mol B, Wilkinson C, Lynch J. Implications of caesarean section for children's school achievement: A population-based study. *The Australian and New Zealand Journal of Obstetrics and Gynaecology*. 2016; 56 (4): 374-380. DOI: 10.1111/ajo.12475
10. Lynch J, Pilkington R, Chittleborough C. Contact with the Child Protection system in South Australia 1999-2015. Adelaide: School of Public Health. The University of Adelaide. 2016.

Suggested citation

Nuske T, Pilkington R, Gialamas A, Chittleborough C, Smithers L, Lynch J. The Early Childhood Data Project. Research Series 2016. Adelaide: School of Public Health, The University of Adelaide, 2016.

References

1. The Australian Government. Australian Institute of Health and Welfare. AIHW. *Child Health, Development and Wellbeing: Risk Factors*. Available from: <http://www.aihw.gov.au/child-health/risk-factors/>
2. Heckman J. Schools, skills and synapses. *Economic Inquiry*. 2008; 46 (3): 289-324.
3. The Marmot Review. Fair Society, health lives. London: UK Department of Health, 2010.
4. National Research Council and Institute of Medicine. From neurons to neighbourhoods: the science of early childhood development. In: Shonkoff J, Phillips D, editors. Washington, DC: National Academy Press; 2000.
5. Sawyer A, Gialamas A, Pearce A, Sawyer M, Lynch J. Five by Five: A Supporting Systems Framework for Child Health and Development. BetterStart Child Health and Development Research Group, School of population Health, The University of Adelaide. 2014.
6. Sen A. Development – Which way now. *Journal of Economics*. 1983; 93 (372):745-62.
7. Shonkoff J, Richter L, van der GAAg J, Bhutta Z. An integrated Scientific Framework for Child Survival and Early Childhood Development. *Pediatrics*. 2012; 129 (2): E460-E72.
8. Council of Australian Governments. Investing in the early years-a national early childhood development strategy. Canberra: COAG; 2009.
9. Government of South Australia. South Australia's Strategic Plan. Adelaide, 2011.
10. Guiver T. Sampling-Based Clerical Review Methods in Probabilistic Sampling. Canberra, ACT: ABS, 2011 1351.0.55.034
11. Holman C, Bass A, Rouse I, Hobbs M. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Australian and New Zealand Journal of Public Health*. 1999; 23(5):453-9.
12. Centre for Health Record Linkage. Quality Assurance. 2012 [updated May 2014]; available from: <http://www.cherel.org.au/quality-assurance>
13. Gialamas A, Pilkington R, Berry J, Scalzi D, Gibson O, Brown A, Lynch J. Identification of Aboriginal children using linked administrative data: Consequences for measuring inequalities. *Journal of Paediatrics and Child Health*. 2016; 52 (5): 534-540.
14. Sawyer M, Pfeiffer S, Sawyer A, Bowering K, Jeffs D, Lynch J. Effectiveness of nurse home visiting for families in rural South Australia. *Journal of Paediatrics and Child Health*. 2014; 50 (12): 1013-1022.

Table 1: Description of data currently available within birth year for South Australian birth cohorts born from the 1st of January, 1999 to the 31st of December, 2013

| Data | Birth Year | | | | | | | | | | | | | | | | Age* |
|--------------------------------------------|------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|------|
| | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | | |
| Perinatal, Congenital Abnormalities | | | | | | | | | | | | | | | | | |
| Births | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| Births Defects Register | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| Neonatal Hearing Screening Program | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| 1-4 week Health Check | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| National Perinatal Depression Initiative | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| Pathways To Parenting | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| Family Home Visit | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| Brief response | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| 6-8 week Health Check | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| 6 month Health Check | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 0 | |
| | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 1 | |
| 18-24 month Health Check | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 1.5 | |
| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | 3 | |
| 4 year Health Check | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | 4 | |
| Public School Enrolment | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | 5 | |
| AEDC | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | 5 | |
| English as an Additional Language/Dialect | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | 6 | |
| Reading Yr1 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | 6 | |
| Reading Yr2 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | 7 | |
| NAPLAN Yr3 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | | 8 | |
| | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | | | 9 | |
| NAPLAN Yr5 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | | | | 10 | |
| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | | | | | 11 | |
| NAPLAN Yr7 | 2011 | 2012 | 2013 | 2014 | 2015 | | | | | | | | | | | 12 | |
| | 2012 | 2013 | 2014 | 2015 | | | | | | | | | | | | 13 | |
| NAPLAN Yr9 | 2013 | 2014 | 2015 | | | | | | | | | | | | | 14 | |
| | 2014 | 2015 | | | | | | | | | | | | | | 15 | |

Available at all ages

| | | | | | | | | | | | | | | | | |
|----------------------------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Hospital Separations (ISAAC) | 2001 | 2001 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Emergency Presentations | 2003 | 2003 | 2003 | 2003 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Child Protection | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Public Housing | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Public Dental: Patient services data | 2000 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Public Dental: Decayed, Missing, Filled Teeth data | 2001 | 2001 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Deaths Registry | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| SA Immunisations | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |

*Age will vary dependant on date of birth and date of data collection. For example the mean age of AEDC is ~5.5 years, while the majority of AEDC data falls within a range of 5 to 7 years of age.