

Appendix 1: Study 3 – requested data items

study variables

Category	Data items/domains
Family health, history and demographics	– Month and year of birth (M/YYYY) – Gestational age at birth – Postcode/suburb at time of each visit (if known) – Language spoken at home – First language (dropdown list) – Main language spoken at home (dropdown list) – Number of languages spoken (total) – Interpreter required (Yes/No) – Type of housing (public, private, town camp) – Household/Family income (if known) – Financial stress (tick box) – Lack of social support/isolation (tick box) – Family worries or concerns (Yes/No collected HU5K child checks) – Recent family stress (tick box) – Family/household food (in)security (if recorded) – Parent/carer education (including highest level of school completed) – Parent/carer marital status – Parent/carer employment status – Family/Carer situation (dropdown list) – Housing conditions (including any urgent housing repairs needed) – Number of children living in the house – Number of siblings – Number of adults living in the house – Number of bedrooms in house – Mother or carer with disability – Other children in family with disability – Siblings with malnutrition or FTT – Family barriers – total (numeric) – Family Hx of respiratory problems (Yes/No) – Family Hx of ear problems (Yes/No) – Family Hx of eye problems (Yes/No) – Parent/Carer mental health issues (yes/no) – Substance abuse in immediate family (Yes/No) – Substance use history taken/updated (tick box) – Parent/Carer mental health issues (yes/no)
Early childhood Education	– Childcare/ Preschool Readiness Program (PRP)/ Preschool enrolment data ○ Days enrolled ○ Months enrolled
Education	– Timestamped records of 3A intensive sessions attended – Timestamped record of Learning and Cultural plans in place – Any recorded reference to primary school performance (e.g., Grade 3 or 5 NAPLAN results, Australian Early Development Census (AEDC) results) – Educational progress (notes)

Continued

Appendix 1: Study 3 – requested data items

study variables

Category	Data items/domains
Health and development	– Birth details including: Mode of birth, duration of labour (numeric), premature birth <37 weeks (tick box), Twin or triplets? Yes/No, APGAR scores, IUGR/Premature/Low Birth Weight (Yes/No), list any birth/neonatal/abnormalities, list of maternal complications, feeding method at birth, history of alcohol or drug use in pregnancy, NT Hospital of birth or place of birth, first time or young mother, parity/gravidity, weeks antenatal care provided, EPDS score (if known). – Baby/Child's Gender – Indigenous status (child's father, child's mother) – Birth weight (grams) – Weight (timestamped) – Height or length including growth centiles (timestamped) – Feeding method (at birth and serial timepoints) – Breastfeeding (exclusively, last 24 hours, at discharge hospital) – Introduction of solids – Diet assessments – Major diagnoses (such as, anaemia, asthma, ear infections etc.) – Pathology results (if known) – Immunisation records (vaccine type/dose/date administered) – Timestamped data/scores and notes for developmental screening/checks (such as ASQ-TRAK) – Timestamped data/scores and notes for developmental assessments, including Paediatric assessments, and formal developmental assessments (such as, The Bayley Scales of Infant and Toddler Development etc.) – Timestamped data/scores and notes for any language assessments – Timestamped data/scores and notes for any educational assessments – Is child exposed to passive smoking? – Does anyone smoke in the house or car? – Developmental flags (drop down list) – Does your child sleep well at night? (Yes/No) – Development: Toilet training (nappies) – Child barriers total (numeric) – Child mood assessment (if recorded) – Parent/child interaction (MEMO) – Medication
Healthcare and service utilisation	– visits to Congress clinics and services (including location of clinic or service, date of visit) – all-cause hospital admissions (including date of admission, length of stay, reason for admission, ICD-10 codes etc.) – all-cause ER presentations (including date of presentation, reason for presentation, ICD-10 codes etc.) – referrals to allied health providers/ CYATS/other specialist services and date/outcome if known – visits to allied health providers/ CYATS/other specialist services and date/outcome if known – Interagency / partner meetings held and date/outcome if known – Utilisation of Congress transport – Child identified at risk (Yes/No) – Child Protection notification made (Yes/No) – Siblings notified to FACS (Yes/No) – Social worker (BoT) assessments (such as scores from Strengths and Difficulties Questionnaire, Strong Souls Assessment Tool, Wechsler – Non-Verbal scale of ability, Wechsler Individual Achievement Test, Adaptative Behavior Assessment System)

Appendix 2: Study 3 - Extracted data items

	Birth	~8 Weeks	4 Months	6 Months	9 Months	12 Months	18 Months	2 Years	2.5 Years	3 Years	3.5 Years	4 Years	4.5 Years	5 Years
Variable name														
Child month/year of birth	☑													
Child sex	☑													
Child Aboriginal status	☑													
Birth weight (kg)	☑													
Birth length (cm)	☑													
Gestational age (weeks)	☑													
NT hospital of birth	☑													
5-minute APGAR score	☑													
Head circumference (cm)	☑													
Breastfeeding at hospital discharge	☑													
Weight (kg)		☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Height (cm)		☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
BMI		☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Feeding method		☑	☑	☑	☑	☑	☑	☑		☑				
Skin assessment		☑	☑	☑	☑	☑	☑	☑		☑		☑		
Recurrent/ Persistent otitis media				☑	☑	☑	☑	☑		☑		☑		
Respiratory assessment				☑	☑	☑	☑	☑		☑		☑		
Dental referral				☑	☑	☑	☑	☑		☑		☑		
Hearing referral				☑	☑	☑	☑	☑		☑		☑		
Other referrals	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
Immunisation status			☑	☑	☑	☑	☑	☑				☑		
Haemoglobin check				☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Iron deficiency treatment	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
Main carer/ Family carer situation*	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
Does anyone smoke in the house or car?		☑	☑	☑	☑	☑	☑	☑		☑		☑		☑
How many adults live in your house?		☑	☑	☑	☑	☑	☑	☑		☑		☑		☑
How many children live in your house?		☑	☑	☑	☑	☑	☑	☑		☑		☑		☑
Housing type	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
How many bedrooms are in the house?		☑	☑	☑	☑	☑	☑	☑		☑		☑		☑
How many languages do you speak?	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
First language	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
Main language spoken at home	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
ASQ-TRAK raw score (Domain 1) communication				☑	☑	☑	☑	☑		☑		☑		
ASQ-TRAK raw score (Domain 2) gross motor				☑	☑	☑	☑	☑		☑		☑		
ASQ-TRAK raw score (Domain 3) fine motor				☑	☑	☑	☑	☑		☑		☑		
ASQ-TRAK raw score (Domain 4) problem solving				☑	☑	☑	☑	☑		☑		☑		
ASQ-TRAK raw score (Domain 5) personal-social				☑	☑	☑	☑	☑		☑		☑		
Number of domains below cut-off				☑	☑	☑	☑	☑		☑		☑		
Referrals made following ASQ-TRAK assessments				↕		↕	↕	↕		↕		↕		
Has the child started preschool? *												☑		

☑ Data items collected as part of routine healthcare.

↕ Data item collected ad hoc/randomly during non-routine or acute presentations.

*Derived composite variable generated by merging 1 or more matched/harmonised variables.



Appendix 3. Study 3 - Detailed data preparation steps

In preparing the dataset for analysis, we undertook the following steps:

- (1) Records of ineligible children were removed (i.e., children not born in 2015, and/or with only a single visit recorded over the study period).
- (2) Duplicate records (i.e., the same variable populated on the same clinical date for the same child) were documented. We cross-checked the values in the duplicate records to see if they were consistent. If continuous values that were found to be inconsistent, we took the smallest value to true. Duplicate records were then dropped.
- (3) Each variable was checked for outliers or implausible values. After documenting the number of outliers or implausible values, we re-coded these as missing.
- (4) We assigned the first day of each month as a proxy “day” for their “date of birth” with Communicare provided “month/year” variable.
- (5) Then we derived a new variable “child age at visit” using new “date of birth” variable and each clinical

date. “Child age at visit” was created to categorise each visit into age-based health assessment milestones. The child’s age at each visit (in years) was rounded to the nearest 0.25 using the STATA command `gen cage = round(clinicalitemage, 0.25)`, enabling visits to be grouped within three-month windows of scheduled milestone ages. For example, if a child attended a visit at 12.2 months, this would be grouped under the 12-month milestone, as it falls within the 10.5 to 13.5-month window. To better reflect observed patterns of attendance, we combined the First Visit and 8-week visit into a single milestone (0–8 weeks).

- (6) Checked derived age at visit for duplicate records (i.e., the same variable populated on the same clinical date for the same child). We cross-checked the values in the duplicate records to see if they were consistent. If continuous values that were found to be inconsistent, we took the smallest value to true. Duplicate records were then dropped.
- (7) We then reshaped the raw data into a format conducive to statistical analysis with 1 row per child and each child’s associated data in columns to the right.

The earliest clinical date per child was considered to be their baseline episode of care date. Multiple records occurring on the same day were defined as one visit.



Appendix 4: Study 3 – Mean, median weight/height/bmi

Table a: Weight, height and BMI in sub-sample of Congress boys (n=9) with complete longitudinal data collected at birth, 6, 12, 18 months, 2, 3 and 4 years of age

Congress boys (n = 9)															
Weight (kg)						Height/Length (cm)					BMI (kg/m ²)				
	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max
Birth	3.18	0.48	3.23	2.20	3.89	50.20	2.39	51.00	47.00	53.50	-	-	-	-	-
6 months	8.11	1.34	7.94	5.49	9.94	66.33	3.07	67.00	59.00	70.50	-	-	-	-	-
12 months	10.10	1.50	10.18	7.50	12.10	75.57	2.61	76.00	71.50	79.50	-	-	-	-	-
18 months	11.52	2.20	10.85	8.61	16.22	81.77	3.37	82.50	77.00	86.00	-	-	-	-	-
2 years	12.96	2.57	12.70	9.96	18.90	85.83	3.53	88.00	79.00	89.00	-	-	-	-	-
3 years	14.16	1.94	14.00	11.60	18.20	92.67	4.39	93.70	84.00	100.00	16.51	2.19	16.41	13.54	21.50
4 years	16.37	2.57	15.9	12.70	21.40	100.11	5.69	101.50	88.00	106.00	16.29	1.66	15.88	13.36	19.59

Table b: Weight, height and BMI in sub-sample of Congress girls (n=9) with complete longitudinal data collected at birth, 6, 12, 18 months, 2, 3 and 4 years of age

Congress girls (n = 9)															
Weight (kg)						Height/Length (cm)					BMI (kg/m ²)				
	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max
Birth	2.99	0.39	2.98	2.53	3.70	50.27	1.69	50.00	48.00	53.50	-	-	-	-	-
6 months	7.23	0.95	7.40	5.75	8.74	65.33	3.36	65.00	61.00	71.00	-	-	-	-	-
12 months	9.26	0.89	9.28	7.57	10.92	74.33	3.67	75.00	66.00	78.00	-	-	-	-	-
18 months	10.50	1.17	10.78	8.53	12.10	79.77	3.11	79.00	75.50	84.00	-	-	-	-	-
2 years	12.07	1.42	12.10	9.10	13.70	84.38	4.02	86.00	75.50	88.00	-	-	-	-	-
3 years	14.63	1.75	14.50	11.60	17.30	93.46	4.41	93.20	86.00	99.30	16.73	1.50	16.63	14.87	20.00
4 years	16.78	1.91	17.00	13.40	19.50	101.11	5.30	101.00	92.00	107.00	16.40	1.31	15.83	14.65	18.42



Appendix 5: Study 3 - Constructs captured across multiple data items

Exact match	Partial match	No match
– Birthweight (SOURCE: Extraction identifier) – Birthweight (kg) SOURCE: (Birth details form) – Baby's birthweight (kg) SOURCE: (low birth weight/premature infant plan) – Gestational age at birth (weeks) SOURCE: (Birth details form) – Gestational age at birth (weeks) SOURCE: (low birth weight/premature infant plan) – Gestational age at birth (weeks) SOURCE: (ANFPP infant health care form) – Number of adults per household (Numeric, discrete) SOURCE: (HU5K child health check) – How many adults live in your house? (Numeric, discrete) SOURCE: (Social and Family History) – Number of children per household (Numeric, discrete) SOURCE: (HU5K child health check) – How many children live in your house? (Numeric, discrete) SOURCE: (HU5K child health check)	– Breastfeeding exclusively Yes/No SOURCE: (HU5K; Child Hx and Risk Assessment) – Breastfeeding at hospital discharge SOURCE: (HU5K; Child Hx and Risk Assessment) – Feeding method (drop down list – “Breast only”) SOURCE: (Birth details form) – Feeding method (drop down list – “Breast only”) SOURCE: (HU5K child health check) – Has the child started preschool? – Enrolled at preschool? – Enrolled at CHAD – Date enrolled in preschool	– First language (drop down list, plus “other” option for free text) SOURCE: (PRP initial assessment) – Main language spoke at home (drop down list, plus “other” option for free text) SOURCE: (Social and Family History) – How many languages do you speak? (Numeric, discrete) SOURCE: (Social and Family History)



Appendix 6. Study 3 – Record statement

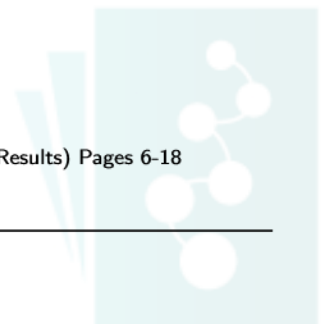
The REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement is a checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

Source: Benchimol EI et al. (2015) The REporting of studies Conducted using Observational Routinely collected health Data (RECORD) Statement. PLoS Med 12(10): e1001885. <https://doi.org/10.1371/journal.pmed.1001885>.

	Item no.	STROBE items	Location in manuscript where items are reported	STROBE items	Location in manuscript where items are reported
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found		RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Title page
Introduction					
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported			(Introduction) Pages 3-4
Objectives	3	State specific objectives, including any prespecified hypotheses			(Introduction) Page 4
Methods					
Study design	4	Present key elements of study design early in the paper			(Methods) Pages 4-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection			(Abstract Introduction and Methods) Pages 2-6
Participants	6	(a) Cohort study Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	(Methods) Pages 4-6 Supplementary Appendix 3



	Item no.	STROBE items	Location in manuscript where items are reported	STROBE items	Location in manuscript where items are reported
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	(Methods) Pages 4-6 Supplementary Appendix 2 and Table 3 (pages 14-18)
Data sources/measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group			(Methods) Pages 4-6 Supplementary Appendices 2, 3 and 5
Bias	9	Describe any efforts to address potential sources of bias			Bias/Limitations discussed on Pages 19-21
Study size	10	Explain how the study size was arrived at			(Methods) Pages 4-6 Supplementary Appendix 3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why			(Methods) Pages 4-6 Supplementary Appendix 3
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study - If applicable, explain how loss to follow-up was addressed			(Methods) Pages 4-6 Supplementary Appendix 3
Data access and cleaning methods				RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	(Methods) Pages 4-6 Supplementary Appendix 3
Results Participants	13	(a) Report the numbers of individuals at each stage of the study (e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (i.e., study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	(Results) Pages 6-18
Descriptive data	14	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) Cohort study - summarise follow-up time (e.g., average and total amount)			(Results) Pages 6-18
Outcome data	15	Cohort study - Report numbers of outcome events or summary measures over time			(Results) Pages 6-18



	Item no.	STROBE items	Location in manuscript where items are reported	STROBE items	Location in manuscript where items are reported
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder -adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorised (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			(Results) Pages 6-18
Other analyses	17	Report other analyses done —e.g., analyses of subgroups and interactions, and sensitivity analyses			(Results) Pages 6-18
Discussion Key results		Summarise key results with reference to study objectives			(Discussion) Pages 19-22
Limitations		Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias		RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	(Discussion) Pages 19-21
Interpretation		Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			(Discussion) Pages 21-22
Generalisability		Discuss the generalisability (external validity) of the study results			(Discussion) Pages 19
Other Information Funding		Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based			Page 22
Accessibility of protocol, raw data and programming code				RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Page 22



