



Measuring vulnerability and disadvantage in early childhood data collections: Phase two

A report for the Australian Government Department of Education

APRIL 2024



Measuring vulnerability and disadvantage in early childhood data collections: A report for the Australian Government Department of Education

Suggested citation: Pham, C., Downes, M., Guo, S., Jahan, F., De Silva, S., O'Connor, E., Gray, S., Priest, N. & Goldfeld, S. (2024). Measuring vulnerability and disadvantage in early childhood data collections: Phase Two. Centre for Community Child Health. Melbourne, Victoria.

<https://doi.org/10.25374/MCRI.26300749>

This report was prepared by the Centre for Community Child Health for the Australian Government Department of Education. This work was undertaken in partnership with the University of Melbourne.

This work is copyright. To request permission to reproduce any part of this material please contact the Centre for Community Child Health.

The Centre for Community Child Health is a department of The Royal Children's Hospital and a research group of the Murdoch Children's Research Institute.

Centre for Community Child Health

The Royal Children's Hospital Melbourne
50 Flemington Road, Parkville
Victoria 3052 Australia

Telephone: +61 9345 6150

Email: enquiries.ccch@rch.org.au

ccch.org.au

The Centre for Community Child Health acknowledges the Traditional Owners of the land on which we work and pay our respect to Elders past, present and emerging.



Contents

Executive summary	6
Background.....	6
Aim of the project	6
Summary of methods.....	7
Key findings.....	7
Recommendations	10
Background	11
Project overview	13
General Methods	15
Data source	15
Measures	15
Statistical analysis plan	17
Part 1 Associations of child disadvantage and priority population indicators with children’s developmental vulnerability on specific domains	18
Overview	18
Methods.....	18
Results.....	19
Conclusion	33
Part 2 Comparison of child-level and area-level indicators of disadvantage	34
Overview	34
Methods.....	34
Results	35
Conclusion	44
Part 3 Child-level disadvantage and priority population indicators for the prediction of children’s developmental vulnerability	45
Overview	45
Methods	45
Results.....	48
Conclusion	53
Conclusion	55
Overall summary	55
Data gaps and challenges	56
Recommendations	57
References	59
Appendices	61
Appendix A. Part 1: Developmental vulnerability on five specific domains	61
Appendix B. Part 2: Child-level and area-level disadvantage indicators	74
Appendix C. Part 3: Best child-level disadvantage predictors of DV1	77



List of figures

Figure E1.	Top 10 child disadvantage and priority population indicators that best predicted DV1	9
Figure 1.	Project workflow for Phase One and Two	14
Figure 2.	Percentage of children’s developmental vulnerability on specific domains by levels of the top two sociodemographic indicators	20
Figure 3.	Percentage of children’s developmental vulnerability on specific domains by levels of the top two health condition indicators	21
Figure 4.	Percentage of children’s developmental vulnerability on specific domains by levels of the top two geographic indicators	22
Figure 5.	Percentage of children’s developmental vulnerability on specific domains by levels of the top two risk factor indicators.	23
Figure 6. 1.	Percentage of children’s developmental vulnerability on specific domains by levels of a top priority population indicator – Child not proficient in English	24
Figure 6.2.	Percentage of children’s developmental vulnerability on specific domains by levels of a top priority population indicator – Parental ancestry	25
Figure 7.	Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the PHY-WELL domain	26
Figure 8.	Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the EMO domain	27
Figure 9.	Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the SOC domain.....	28
Figure 10.	Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the LANG-COG domain	29
Figure 11.	Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the COM-KNO domain.....	30
Figure 12.	Association between SEIFA quintile groups and DV1.....	43
Figure 13.	An overview of the methodology of Part 3 in this Phase Two report.....	46
Figure 14.	The 10 child-level disadvantage and priority population indicators in candidate Model B	53
Figure C.1.	Correlation matrix for child-level disadvantage indicators in the sociodemographic lens.	79
Figure C.2.	Correlation matrix for child-level disadvantage indicators in the health conditions lens.....	81
Figure C.3.	Correlation matrix for child-level disadvantage indicators in the geographic lens.	83
Figure C.4.	Correlation matrix for child disadvantage indicators in the risk factors lens	85
Figure C.5.	Correlation matrix for priority population indicators.....	87



List of tables

Table E1.	A summary of the common disadvantage and priority population indicators related to children's developmental vulnerability across the five specific developmental domains.....	8
Table 1.	Summary of the 36 key child-level disadvantage indicators and 1 area-level disadvantage indicator across 4 disadvantage lenses and 7 priority population indicators examined in this project.....	16
Table 2.	AEDC domain descriptions.....	17
Table 3.	Univariable associations between the priority population indicators and children's developmental vulnerability on specific developmental domains.....	31
Table 4.	Summary of the child disadvantage and priority population indicators identified in the top 15 shortlist for developmental vulnerability on one or more specific domains.....	32
Table 5.	Cross tabulations between child-level and area-level disadvantage indicators.....	36
Table 6.	Summary of the 17 candidate child disadvantage and priority population indicators as predictors of DV1 ordered by magnitude of univariable association.....	49
Table 7.	Summary of the child disadvantage and priority population indicators in the four candidate models for the prediction of DV1.....	51
Table 8.	Predictive performance of the four candidate models using 60% training and 40% test data (2-fold cross validation).....	52
Table 9.	Predictive performance of SuperLearner individual learners and ensemble algorithm for the best candidate model.....	53
Table A.1.	The percentage of children's developmental vulnerability on five specific domains by the key child disadvantage indicators.....	61
Table A.2.	Associations between child disadvantage indicators across the four social determinant lenses and developmental vulnerability on five specific domains.....	68
Table A.3.	The percentage of children's developmental vulnerability on five specific domains by the priority population indicators.....	72
Table B.1.	Distribution of DV1 by area-level child disadvantage (SEIFA).....	74
Table B.2.	Associations of area-level and child-level disadvantage indicators with children's DV1.....	74
Table C.1.	Summary of child-level disadvantage indicators in the sociodemographic lens as predictors of DV1 ordered by magnitude of univariable association.....	77
Table C.2.	Summary of child disadvantage indicators in the health conditions lens as predictors of DV1 ordered by magnitude of univariable association.....	80
Table C.3.	Summary of child disadvantage indicators in the geographic lens as predictors of DV1 ordered by magnitude of univariable association.....	82
Table C.4.	Summary of child disadvantage indicators in the risk factor lens as predictors of DV1 ordered by magnitude of univariable association.....	84



Executive summary

Background

Early childhood experiences and environments affect children's health, development and wellbeing throughout their lifetime. Children may experience disadvantage depending on the conditions in which they live, learn and grow. This can lead to immediate and long-term impacts at both individual and societal levels. The COVID-19 pandemic has likely increased existing disadvantage for these children, contributing to worsened outcomes and increased inequity. Robust measurement of disadvantage during early childhood is essential for identifying effective strategies to address these inequities to optimise children's health, development and wellbeing.

The Multi-Agency Data Integration Project (MADIP) First Five Years (FFY) project is an Australian Government administrative dataset that includes the Australian Early Development Census (AEDC). The AEDC is a valuable tool for monitoring childhood developmental inequities, assessing aspects of children's development across five key domains in the first year of full-time school. Together, this data can help to build understanding of the effects of multidimensional early childhood factors on children's health, development and wellbeing, and identify children at higher risk of developmental vulnerability.

In our Phase One work, we used the MADIP-FFY-AEDC data, in collaboration with the Australian Government Department of Education, to conduct a rapid desktop review and data evaluation that demonstrated a range of factors that drove inequitable developmental outcomes in children. Our current Phase Two work expands on this work to further understand associations between key child disadvantage and priority population indicators and childhood developmental vulnerability. The key child disadvantage indicators in this project are guided by the Changing Children's Chances (CCC) social determinants framework. Phase Two findings will provide further valuable insights into the subset of disadvantage and priority population indicators that best predict children's developmental vulnerability that could be leveraged for policy purposes.

Aim of the project

This Phase Two project builds on our Phase One work and seeks to further understand the associations between key child-level and area-level disadvantage and priority population indicators and children's developmental vulnerability, by addressing the following three aims:

1. To investigate associations of child disadvantage and priority population indicators with children's developmental vulnerability on the **five specific developmental domains** (physical health and wellbeing, social competence, emotional maturity, language and cognitive skills (school-based), and communication skills and general knowledge).
2. To compare child-level and **area-level measures of disadvantage**.
3. To determine the **best combination** of child-level disadvantage and priority population indicators for identifying children who are at the highest risk of developmental vulnerability on one or more domain(s) (DV1).



Summary of methods

Similar to our Phase One work, we used the MADIP-FFY dataset, including the AEDC developmental outcomes, to identify, evaluate and analyse 36 child disadvantage indicators. These indicators were guided by the CCC social determinants framework, which is structured around four social determinant lenses (i.e., sociodemographic, health conditions, geographic and risk factors; referred to as disadvantage lenses hereafter). We also examined the priority population indicators related to children's developmental vulnerability outcomes.

In this Phase Two project, additional data analysis was conducted in three parts:

- **Part 1:** Associations of child disadvantage and priority population indicators with children's developmental vulnerability on five specific developmental domains. We used descriptive statistics to understand the distribution of each of the child disadvantage and priority population indicators and the five developmental domains. Separate univariable regression analyses were conducted to estimate the association of each child-level disadvantage and priority population indicator with children's developmental vulnerability on each specific developmental domain.
- **Part 2:** Comparison of child-level and area-level measures of disadvantage. We used cross-tabulations to summarise the distribution of each of the 36 child-level disadvantage indicators by area-level Socio-Economic Indexes for Areas (SEIFA) quintiles. We also estimated the univariable association of the area-level disadvantage measure with children's DV1, and then compared the magnitude of associations with child-level disadvantage indicators from our Phase One work.
- **Part 3:** The best combination of child-level disadvantage and priority population indicators that best predicted DV1. We used logistic regression analyses and machine learning approaches to identify an optimal set of predictors of children's DV1.

Key findings

Part 1: Associations of child disadvantage and priority population indicators with children's developmental vulnerability on specific developmental domains

Univariable regression analyses were conducted to estimate the association of each child disadvantage and priority population indicator with children's developmental vulnerability on five domains. The ranking of indicators by magnitude of univariable association with developmental vulnerability for each domain was broadly consistent with the ranking for DV1 in the Phase One report.

Key domains of children's developmental vulnerability

Physical health and wellbeing (PHY-WELL)

- Children's physical readiness for the school day, physical independence and gross and fine motor skills.

Social competence (SOC)

- Children's overall social competence, responsibility and respect, approach to learning and readiness to explore new things.

Emotional maturity (EMO)

- Children's pro-social and helping behaviours and absence of anxious and fearful behaviour, aggressive behaviour and hyperactivity and inattention.



Language and cognitive skills (school-based) (LANG-COG)

- Children's basic literacy, advanced literacy, basic numeracy, and interest in literacy, numeracy and memory.

Communication skills and general knowledge (COM-KNO)

- Children's communication skills and general knowledge based on broad developmental competencies and skills.

Based on the strength of the univariable associations between the child disadvantage and priority population indicators with children's developmental vulnerability on specific developmental domains, there was considerable overlap between the domains. There were 18 common indicators across these five developmental domains (physical health and wellbeing, social competence, emotional maturity, language and cognitive skills (school-based), and communication skills and general knowledge), as shown in Table E1.

Table E1. A summary of the common disadvantage and priority population indicators related to children's developmental vulnerability across the five specific developmental domains

Sociodemographic (n=7)	Health conditions (n=3)	Geographic (n=2)	Risk factors (n=5)	Priority population (n=1)
<ul style="list-style-type: none"> • Lower income household • Lower parental highest education level • Parent was employed for four years or less • Child lived in a lone parent family • Parent received any type of special childcare benefit • Parent received any type of social support payment • Parent highest occupation 	<ul style="list-style-type: none"> • Child has had any chronic health issue(s) • Child has had any mental health issue(s) • Parent has had any mental health issue(s) 	<ul style="list-style-type: none"> • House crowding (3 or more additional bedrooms needed) • Rented tenure type 	<ul style="list-style-type: none"> • Child is not regularly read to at home • Child was born to a teenage mother • Child did not attend preschool • Child experienced the death of a parent • Child had no unpaid childcare 	<ul style="list-style-type: none"> • Child not proficient in English

Indicator 'child not regularly read to at home' had the strongest association with children's developmental vulnerability on all five domains. Indicator 'Child Care Subsidy income thresholds' was among the top three indicators on children's developmental vulnerability on all five domains, together with 'parental highest education level' on the PHY-WELL, LANG-COG and COM-KNO domains and 'child has had any mental health issues' on the EMO and SOC domains.

Part 2: Comparison of child-level and area-level measures of disadvantage

We used SEIFA as a proxy for area-level measures of socioeconomic status. We observed clear differences in the distribution of each child-level disadvantage indicator across the SEIFA quintiles. There were higher proportions of children exposed to child-level disadvantage among those living in the most disadvantaged communities than those living in least disadvantaged communities, suggesting that children living in more disadvantaged communities are likely to be at higher risk of experiencing child-level disadvantage. However, for children living in the least disadvantaged communities, 78.6% had family income in the highest bracket of \$99,864 or more while 10.8% had family income in the lowest bracket of \$56,137 or less. These 10.8% of children were considered least disadvantaged based on the area-level measure but disadvantaged on the child-level income measure.



Further, we examined the associations between area-level SEIFA and DV1, and then compared the magnitude of associations with the child-level disadvantage indicators. We found that two child-level indicators i.e., 'lower household income', and 'child not regularly read to at home', had stronger associations with DV1 than area-level SEIFA. Given that there was considerable variation within SEIFA quintiles, area-level measures may not capture the complexity or distribution of disadvantage that children may experience.

Part 3: Child disadvantage and priority population indicators for the prediction of DV1

A subset of the 36 child-level disadvantage indicators and seven priority population indicators identified from Phase One were selected based on the strength of estimated univariable associations with DV1 and the percentage of missing observations. Four candidate prediction models were established, each with a different selection of indicators across the four disadvantage lenses and priority population groups, to identify the best combination of indicators for predicting the risk of children's DV1.

Following consideration of various combinations of disadvantage and priority population indicators, we found that Model B comprising ten predictors in total – two indicators from each of the four disadvantage lenses and two indicators representing priority populations – was the simplest model for predicting DV1 (Figure E1).

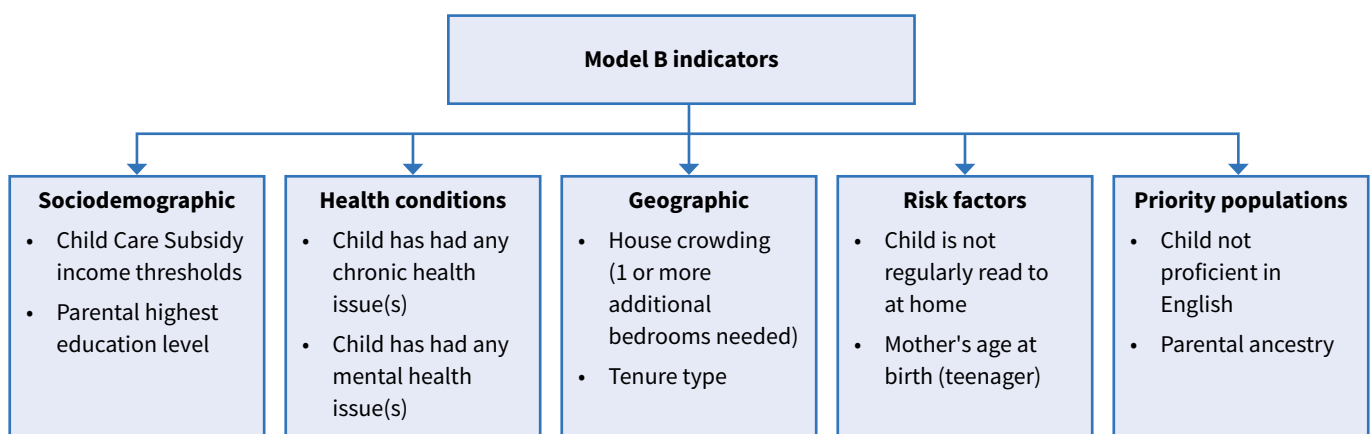


Figure E1. Top 10 child disadvantage and priority population indicators that best predicted DV1

However, this model did not perform particularly well at discriminating between children who were considered developmentally vulnerable overall from those who were not. The overall predictive accuracy of Model B was 85%, indicating that developmental vulnerability status was correctly classified for 85% of children in the test dataset. This was largely driven by a very high specificity of 98%, indicating that the model correctly identified 98% of children who were known or observed to be **not** developmentally vulnerable.

Conversely, the model was much less successful at accurately identifying children who were developmentally vulnerable, doing so for only 27% of those children known or observed to be developmentally vulnerable. The remaining 73% of children observed as developmentally vulnerable were incorrectly classified by the prediction model as not developmentally vulnerable. This indicates that there are likely additional important aspects of a child and their environment, beyond the measured set of disadvantage indicators, that influence the risk of DV1.



Recommendations

The following steps are recommended for the advancement of childhood data collections in future work.

Research

- Examine the distribution of child-level disadvantage indicators and their relationships with developmental vulnerability in these **subgroup populations**.
- Investigate the complex interplay between child-level disadvantage indicators and priority population indicators together on children's developmental vulnerability.
- Investigate the **longitudinal trajectories** of child-level disadvantage and children's developmental vulnerability over time.
- Examine and compare how a model with multiple indicators **within a lens** (e.g., sociodemographic) would perform relative to models with indicators across each of the lenses to predict children who are at the highest risk of developmental vulnerability.
- Further understand the key drivers/mechanisms of developmental inequities through **causal modelling** that can be used for more precise policy interventions.

Data and practice

- Explore **additional geographic indicators** beyond the household level, and carefully consider other housing variables that may assist with the measurement of important key measures such as housing affordability and housing stress.
- To routinely collect 'child ancestry' in future data collections and include measures that accurately capture **diverse backgrounds**.
- Utilise recently linked data with AEDC to indicate a more proximal source of **parental highest education**.
- Carefully consider **other sociodemographic factor** such as single-income families, extended families, and lack of provision of paid childcare.

Policy

- Interpret child-level and area-level findings with caution and acknowledge the **measurement limitations**.
- Governments need to use language that does **not stigmatise** these children when reporting on a negative outcome. For example, 'children experiencing disadvantage' rather than 'disadvantaged children'.
- Addressing childhood disadvantage should consider a **social determinants framework** to better capture the **complexity** of disadvantage.
- The MADIP-FFY data is a **powerful resource**. Governments should consider how this data could be better used to **inform policy and decision making**.



Background

A child's early experiences and environments shape their health, development and wellbeing for life.^{1,2} Experiencing disadvantage during early childhood has lasting impacts at both the individual and societal level; while the impacts can be immediate and/or long-lasting,³ they are proportionate to the level of disadvantage experienced. Disadvantage is both complex and multi-faceted. It depends on the conditions in which children live, learn and grow (social determinants).⁴ Disadvantage can be conceptualised as the aggregation of an individual's health, access to education, employment, the extent to which material basics are met, quality of housing and a range of other factors. The COVID-19 pandemic is likely to have exacerbated existing disadvantage and vulnerability for children, leading to increased inequity and worsened outcomes.

To identify effective strategies to address these inequities,⁶⁻⁸ accurate measurement of the levels of disadvantage experienced by children and their families is necessary.⁸ To ensure that resources are allocated to children who are most vulnerable, policymakers must be able to distinguish groups of children experiencing comparatively higher levels of disadvantage. Given the potential influence of disadvantage on children's development decreases with age,^{9,10} robust measurement of disadvantage experienced during the early childhood years is critical to understanding the breadth of the problem, assessing time-related changes, and identifying interventions to optimise children's health, development and wellbeing trajectories.

The MADIP-FFY data includes the AEDC developmental vulnerability outcomes. The AEDC is a nationwide data collection that is undertaken every three years, first conducted in 2009. By measuring how children are faring in the first year of full-time school, the AEDC aids in understanding which early childhood factors support or hinder children's health, development and wellbeing. Subsequently, the AEDC helps identify areas for improvement to ensure that all children have the best start in life. The AEDC is therefore a valuable tool for understanding and monitoring developmental inequities in childhood. Children's experiences of disadvantage are complex, and the factors contributing to these experiences are also complex, including individual factors (such as children's participation in early childhood education and care (ECEC)), factors related to the child's family's circumstances, and community-level factors. The ability of the MADIP-FFY to capture these multi-dimensional factors may provide a more in-depth measure of children's experiences of inequity and aid the identification of priority populations of children who are at a higher risk of developmental vulnerability.

Our Phase One project¹¹ commenced in 2021, and utilised the MADIP-FFY data in collaboration with the Australian Government Department of Education (referred to as the Department, hereafter). The overall aim of the project was to identify options for feasible child-level indicators of disadvantage. This project was guided by our previous work with the use of our CCC social determinants framework.⁸ In brief, the framework is a robust tool for understanding the multi-dimensional drivers of child inequities⁸ and is structured around four interrelated social determinant 'lenses' (referred to as disadvantage lenses, hereafter). The lenses include:

1. sociodemographic (characteristics that define subpopulation groups)
2. health conditions (diagnosable medical conditions for parents/carers and children)
3. geographic (characteristics of the places where children live)
4. risk factors (attributes, characteristics and exposures that increase the likelihood of poor child outcomes).⁸

The framework captures broader upstream (i.e. wider social, policy and cultural contexts) and downstream (i.e. factors that immediately impact children's lives) social determinants. It is useful for considering pathways and mechanisms and identifying modifiable policy levers to understand and address child inequities.



The main findings of the **Phase One project** included the following:

1. A rapid desktop review showed Australian and State and Territory government agencies are currently drawing on a wide range of data sources and indicators.
 - These indicators are used to report on the factors that shape children's early health, development and wellbeing, and drive inequitable developmental outcomes.
2. A large list of 87 child-level disadvantage indicators were mapped against the MADIP-FFY data to determine availability. A total of 37 available indicators were assessed against the criteria of simplicity, quality and relevance, and further examined.
 - This resulted in identification of the top 15 disadvantage indicators that had the strongest univariable associations with children's developmental vulnerability outcomes.

Our work provided insights into the multifaceted aspects of disadvantage that were most associated with the highest risk of developmental vulnerability in children. This data could be used to inform more **precise policy decisions** to redress child inequities and maximise impact.⁷

The current Phase Two project extends on our previous Phase One work.¹¹ Findings from Phase Two will provide more precise and robust insights on the subset of child-level disadvantage indicators that best predict children's developmental vulnerability. Additionally it provides insights into specific domains of childhood developmental vulnerability and area-level versus child-level measures of disadvantage.



Project overview

This Phase Two project extends the Phase One work¹¹ that provided a detailed investigation on disadvantage and vulnerability indicators within early childhood data collections. The overall aim of this project is to identify options for feasible child-based indicators of disadvantage for the Department's consideration. In this work, the term 'child-level disadvantage indicators' refer to a measurable piece of information relating to aspects of children's experiences of disadvantage or vulnerability.

The key child-level disadvantage indicators in this project are guided by the CCC social determinants framework⁸ which demonstrates that children's development is strongly influenced by the circumstances in which they live, learn and grow. Here, the disadvantage indicators considered manifest across the social determinants at the individual, family and community levels.

Specifically in Phase Two, the Department seeks to further understand the associations between key child-level disadvantage indicators and childhood developmental vulnerability by addressing the following aims:

- To estimate univariable associations of child-level disadvantage indicators with developmental vulnerability on each of the **five specific key domains of childhood developmental vulnerability** (physical health and wellbeing, social competence, emotional maturity, language and cognitive skills (school-based), and communication skills and general knowledge).
- To compare child-level and **area-level indicators of disadvantage** and to estimate univariable associations of area-level measures of disadvantage with DV1 (for comparison to results for child-level disadvantage indicators from Phase One).
- To build upon the univariable analyses completed in Phase One to determine the **best combination** of child-level disadvantage indicators for identifying children most at risk of developmental vulnerability.

An overview of the methodology for the overall project is shown in Figure 1. It includes:

1. Eighty-seven child-level disadvantage indicators were identified in Phase One.¹¹ Based on availability and evaluation, 37 key disadvantage indicators were analysed. These were further examined in this Phase Two project.
2. Additional data analyses were divided into three parts:
 - **Part 1:** Associations of child-level disadvantage and priority population indicators with children's developmental vulnerability on specific developmental domains.
 - **Part 2:** Comparison of child-level and area-level indicators of disadvantage.
 - **Part 3:** Child-level disadvantage and priority population indicators for the prediction of children's developmental vulnerability.
3. The provision of final recommendations.

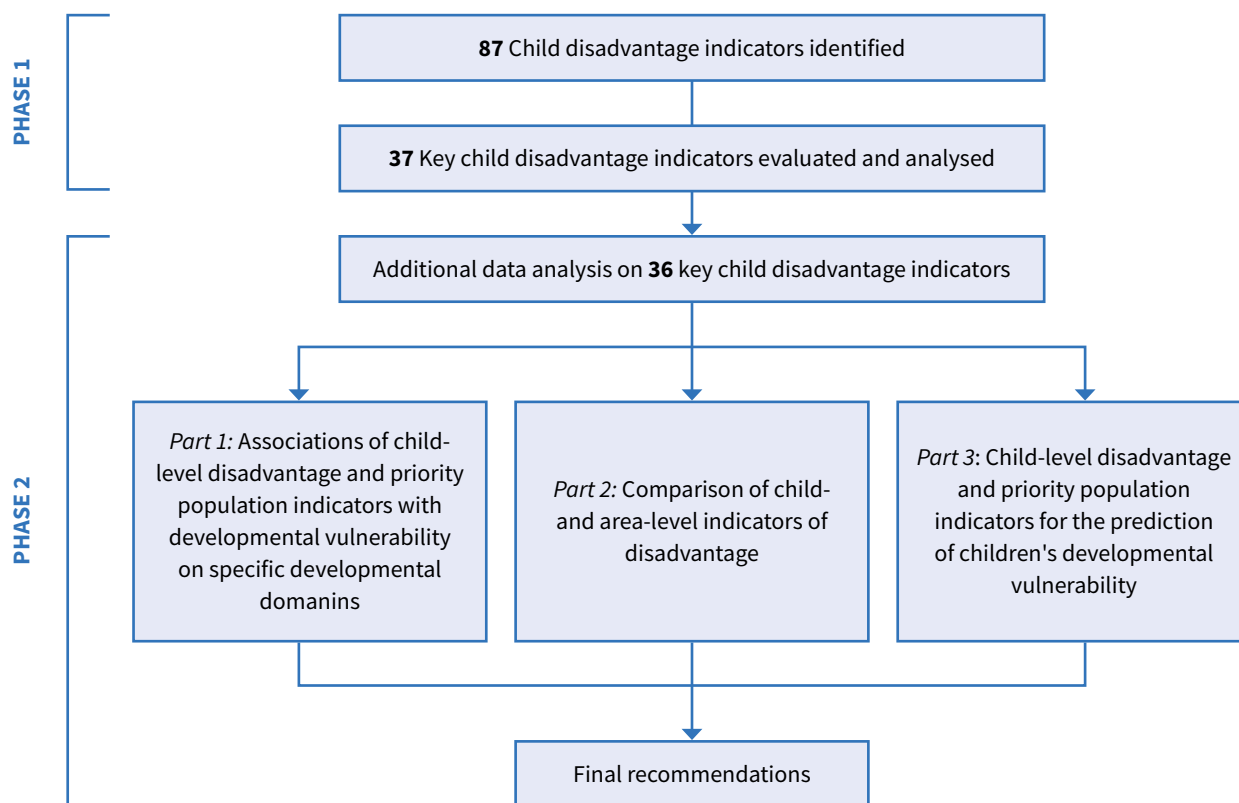


Figure 1. Project workflow for Phase One and Two

Note. 36 of the original 37 child disadvantage indicators were examined in Phase Two; maternal education level was removed due to a high level of missing data.



General Methods

Data source

We continued to draw on deidentified data from the MADIP-FFY project.¹¹ In total, there were 293,910 children with 2018 AEDC outcomes linked to FFY relevant datasets. In brief, the FFY data provide comprehensive information on health, education, economic, social services and population demographics over a child's life from birth to school entry. The FFY data includes linkage with the following datasets:

- Australian Early Development Census (AEDC)
- Census of Population and Housing (Census)
- Child Care Management System (CCMS)
- Data exchange (DEX)
- Data Over Multiple Individual Occurrences (DOMINO)
- Medicare Benefits Schedule (MBS)
- National Health Survey (NHS)
- National Quality Standards (NQS)
- Pay As You Go (PAYG)
- Personal Income Tax (PIT)
- Pharmaceutical Benefits Scheme (PBS)
- Registries of Deaths

Measures

A summary of all the child-level and area-level disadvantage and priority population indicators examined in this report are shown in Table 1. Of note, while we focused on 'child-level' disadvantage, all data were de-identified and there was no risk of reidentification.

Child-level disadvantage indicators

In Phase Two, we used the same 37 child-level indicators from our Phase One report.¹¹ However, only 36 child-level disadvantage indicators (referred to as key child disadvantage indicators, hereafter) were examined in Phase Two due to the removal of maternal education level, which had a high level of missing data. For further details on the measurement and coding of each child-level disadvantage indicator, see the Phase One Report.¹¹

Area-level disadvantage indicator

In Phase Two, we added a new measure of area-level childhood disadvantage, which was assessed by the Socio-Economic Indexes for Areas (SEIFA). Compared to child-level measures of disadvantage (e.g., household income, parental education or occupation), SEIFA is a relative measure, which means that a given geographic area may be classified as disadvantaged relatively rather than disadvantaged in absolute terms.¹²



Priority population indicators

Similarly, the same priority population indicators were used in our Phase Two work. In addition, we added a new priority population indicator (parental ancestry) in this report. Parental ancestry was derived using data from the 2016 Census. Similar to our previous work by Priest et al.,¹³ we used a prioritisation method to create mutually exclusive categories allocating each parent to the category with the highest level of stigmatisation, in the following order (Aboriginal and/or Torres Strait Islander, Pacific Islander/Māori, Middle Eastern, African, Latin American, South Central Asian, Northeast Asian, Southeast Asian, European, and Anglo-Celtic). Due to limited data available from the Census dataset, we included those who were self-identified as 'Australian' in the group 'Anglo-Celtic'. It is important to note that Australian is a nationality, rather than an ancestry or ethnicity. While 'Australian' is often used as a proxy for being white/Anglo-Celtic, this group is exclusionary of those who are self-identified as Australian and also from other backgrounds. Of note, those born overseas in a non-European or English-speaking country and/or with poor proficiency in English are proxies for ethnic minority backgrounds.

Table 1. Summary of the 36 key child-level disadvantage indicators and 1 area-level disadvantage indicator across 4 disadvantage lenses and 7 priority population indicators examined in this project

Sociodemographic (n=17)	Health conditions (n=6)	Geographic (n=5)	Risk factors (n=9)	Priority population (n=7)
<ul style="list-style-type: none"> Household income Equivalised income Poverty line Family eligible for a Low Income Health Care Card Family Tax Benefit A Family Tax Benefit B Child Care Subsidy income thresholds Parental highest education level Parent occupation (3 categories) Parent occupation (2 categories) Parent employment status Parent employment average duration Parent received social support payment Parent received special childcare benefit Household size Child with a lone parent family SEIFA * 	<ul style="list-style-type: none"> Parent has had any chronic health issue(s) Child has had any chronic health issue(s) Parent has had any mental health issue(s) Child has had any mental health issue(s) Parent mental health issue duration Child mental health issue duration 	<ul style="list-style-type: none"> House crowding (3 or more additional bedrooms needed) House crowding (1 or more additional bedrooms needed) Dwelling type Rented tenure type Child has moved residence address in the last 5 years 	<ul style="list-style-type: none"> Preschool non-attendance Childcare non-attendance Average weekly childcare hours Unpaid childcare Child's age-group at childcare entry Child is not regularly read to at home Mother's age at birth (teenager) Mother's age at birth (later) Parental death 	<ul style="list-style-type: none"> Child's Aboriginal and Torres Strait Islander status Child's country of birth Parents' country of birth Child's LBOTE Parents' LBOTE Child not proficient in English Parental ancestry

Asterisk (*) indicate area-level disadvantage indicator. LBOTE language background other than English; SEIFA socio-economic indexes for areas.



Child developmental vulnerability outcomes

As in Phase One,¹¹ children's developmental outcomes were measured using the 2018 AEDC data collection for efficiency and consistency. In brief, the AEDC is a cross-sectional population census designed to record children's development in the first year of full-time school across Australia.¹⁴ The AEDC provides information about children's demographic characteristics (e.g., sex, age, parent education) and five domains of early developmental outcomes.^{15,16}

Children are categorised as developmentally vulnerable on a specific domain if their scores fall below the cut-off score established in 2009, corresponding to the 10th percentile on that domain in the 2009 AEDC data collection.¹⁷ For each of the five AEDC domains, children receive a score between 0 and 10, where 0 is most developmentally vulnerable. The cut-off scores set in 2009 provide a reference point against which later AEDC results can be compared. These have remained the same across all collection cycles. In the first data collection cycle, a series of cut-off scores was established for each of the five domains: children falling below the 10th percentile were categorised as 'developmentally vulnerable'; children falling between the 10th and 25th percentile were categorised as 'developmentally at risk'; and all other children were categorised as 'developmentally on track'.

For this report, analyses focused on one AEDC summary outcome across all the domains, i.e., children who were developmentally vulnerable on one or more domain(s) (referred to as DV1 hereafter). Further, developmental vulnerability on the five specific domains of physical health and wellbeing (PHY-WELL), social competence (SOC), emotional maturity (EMO), language and cognitive skills (school-based) (LANG-COG), and communication skills and general knowledge (COM-KNO) were also examined individually. Detailed summaries of AEDC outcome measures are described in the Phase One Report (see Table 2 below).¹¹

Table 2. AEDC domain descriptions

Physical health and wellbeing (PHY-WELL)	Social competence (SOC)	Emotional maturity (EMO)	Language and cognitive skills (school-based) (LANG-COG)	Communication skills and general knowledge (COM-KNO)
Children's physical readiness for the school day, physical independence and gross and fine motor skills	Children's overall social competence, responsibility and respect, approach to learning and readiness to explore new things	Children's pro-social and helping behaviours and absence of anxious and fearful behaviour, aggressive behaviour and hyperactivity and inattention	Children's basic literacy, advanced literacy, basic numeracy, and interest in literacy, numeracy and memory	Children's communication skills and general knowledge based on broad developmental competencies and skills

Statistical analysis plan

This Phase Two project aimed to deliver three parts of additional analyses:

- **Part 1** included estimation of univariable associations between child-level disadvantage indicators and developmental vulnerability on each of the five developmental domains.
- **Part 2** included comparison of the distribution of previously identified child-level disadvantage indicators across quintiles of the SEIFA area-level disadvantage measure and estimation of the univariable association of area-level disadvantage with children's DV1.
- **Part 3** included examination of the best combination of child-disadvantage indicators for predicting DV1 and was completed over three steps: (1) predictor set selection; (2) identification of an optimal set of predictors; and (3) exploration of ensemble machine learning approaches.

Further details are summarised in the subsequent sections of this report.



Part 1

Associations of child disadvantage and priority population indicators with children's developmental vulnerability on specific domains

Overview

Different child disadvantage indicators have been shown to be associated with specific developmental outcomes in children. The aim of Part 1 was to further understand the associations between the 36 key child disadvantage and seven priority population indicators with children's developmental vulnerability on the five specific domains at school entry. Overall, we found that there was a diverse range of indicators across the four disadvantage lenses and the priority population groups, particularly for the sociodemographic lens, that was associated with children's developmental vulnerability on specific domains.

Methods

Measures

The 36 child-level disadvantage indicators across the four disadvantage lenses and seven priority population indicators were examined in this part to maintain consistency with our Phase One work.¹¹ Developmental vulnerability on each of the five specific domains including physical health and wellbeing (PHY-WELL), social competence (SOC), emotional maturity (EMO), language and cognitive skills (school-based) (LANG-COG), and communication skills and general knowledge (COM-KNO) were considered as separate outcomes. For further details, please refer to the [General Methods](#) section.

Statistical analysis

We first used descriptive statistics to understand the distribution of each of the five developmental domains by each child disadvantage and priority population indicator. We conducted separate univariable regression analyses to estimate the association of each child disadvantage and priority population indicator with children's developmental vulnerability on each specific domain. Estimates were expressed as risk ratios (RR) with 95% confidence intervals (CI), which can be interpreted as the risk of developmental vulnerability on a specific domain among children experiencing disadvantage relative to their peers who were not experiencing disadvantage (RR>1 indicative of increased risk, RR<1 indicative of reduced risk). Confidence intervals describe the uncertainty surrounding of an estimate and can be interpreted as a range of plausible values for the quantity of interest in the target population given the observed data.

For each developmental domain, child disadvantage indicators were ranked according to the strength of their estimated univariable associations (RRs further away from 1 indicative of a stronger association) and the top 15 indicators were presented. Similarly to Phase One, if multiple indicators measured within the same construct (e.g., different ways of categorising household income i.e., poverty line, Family Tax Benefit A and B) were in the shortlisted top 15 indicators for a given developmentally vulnerable domain, only the indicator with the strongest magnitude of univariable association for



that construct was selected. Given that there were limited numbers of priority population indicators, all associations were presented, and these were not ranked.

Results

Key findings

- Children who experienced disadvantage had higher risk of developmental vulnerability on all five domains of developmental vulnerability than those not experiencing disadvantage.
- There were 18 common child-level disadvantage and priority population indicators across the five developmental domains.
- The risk factor indicator ‘child not regularly read to at home’ had the strongest association with developmental vulnerability on all five domains.
- Other indicators with strong associations with developmental vulnerability were ‘Child Care Subsidy income thresholds’, ‘parental highest education’ and ‘child has had any mental health issues’.

Participant characteristics

Overall, the percentage of children developmentally vulnerable on each domain was: 9.4% (PHY-WELL), 8.3% (EMO), 9.5% (SOC), 6.2% (LANG-COG), and 7.5% (COM-KNO). Within each disadvantage lens and priority population group, Figures 2-6.2 showed the proportion of children developmentally vulnerable on each of the five developmental domains by two child disadvantage indicators. The two indicators within each lens that were included in these figures were the ones with the strongest univariable associations with DV1 from our previous Phase One work. The full set of summary statistics including the proportion of children developmentally vulnerable on each of the five specific domains by each of the 36 child disadvantage indicators are shown in [Appendix A, Table A.1](#) and by each of the seven priority population indicators in [Appendix A, Table A.3](#).

Sociodemographic

The top two indicators within the sociodemographic lens were ‘Child Care Subsidy income thresholds with four categories’ and ‘parental highest education level’, which was based on the univariable associations with children’s DV1 in Phase One (Figure 2). Overall, the proportion of children who were developmentally vulnerable on each domain was higher among those who came from families with lower income and had parents who were less educated.

Among children who came from families in the lowest income category, the highest proportion of children who were developmentally vulnerable was on the PHY-WELL domain, followed by SOC and COM-KNO domains. The lowest proportion of children was on the LANG-COG and EMO domains. Among children who came from families in the highest income category, the proportion of children who were developmentally vulnerable was highest on the SOC domain and lowest on the LANG-COG domain.

Among children who had parents with the lowest education level, the highest proportion of children who were developmentally vulnerable was on the PHY-WELL domain, followed by COM-KNO, SOC, and LANG-COG domains. The lowest proportion of children who were developmentally vulnerable was on EMO domain. Among children who had parents with the highest education level, the proportion of children who were developmentally vulnerable was highest on the SOC domain and lowest on the LANG-COG domain.

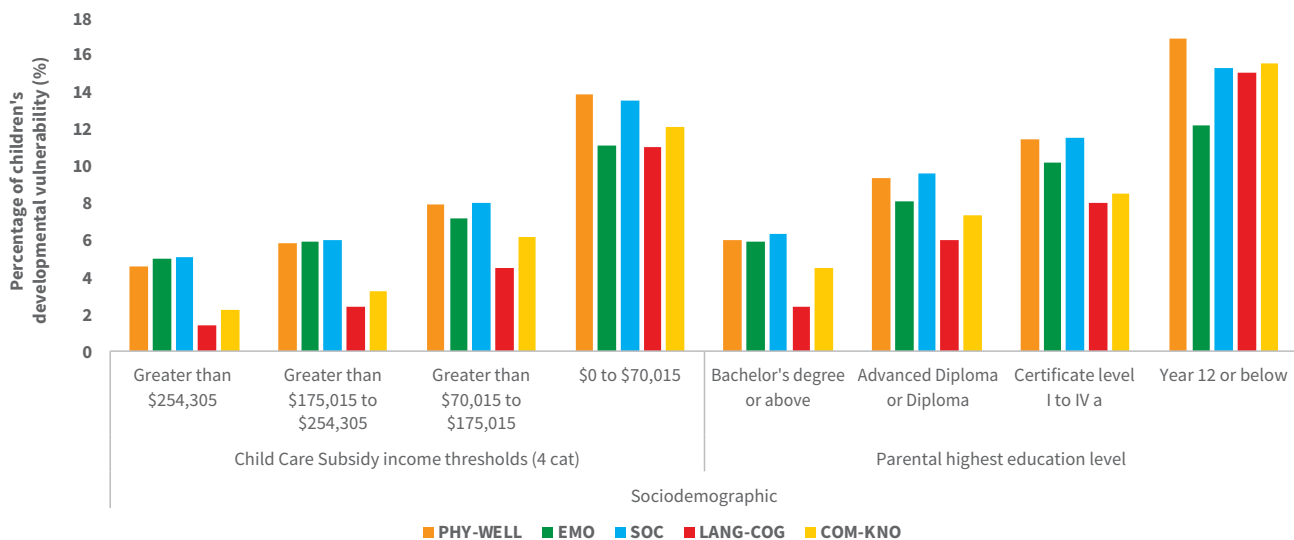


Figure 2. Percentage of children's developmental vulnerability on specific domains by levels of the top two sociodemographic indicators

^a Certificate level I to IV including trade qualification. Note. Top two indicators are based on the magnitude of association per disadvantage lens with children's DV1 from our Phase One work. *Abbreviations:* COM-KNO communication skills and general knowledge; DV1 developmental vulnerability on one or more domain(s); EMO emotional maturity; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.

Health conditions

The top two indicators within the health condition lens were 'child has had any chronic health issue(s)' and 'child has had any mental health issue(s)', which was based on the univariable associations with children's DV1 in Phase One (Figure 3). Of note, these indicators were derived from records from the MBS/PBS, rather than deriving them from teacher-reported data. Overall, the proportion of children who were developmentally vulnerable on each domain was higher among those who had a chronic health issue or mental health issue at any time point from their birth to the time they started school.

Among children with any chronic health issues, consistent with previous research,^{18,19} the highest proportion of children who were developmentally vulnerable was on the SOC domain, followed by the PHY-WELL, EMO, and COM-KNO domains. The lowest proportion was on the LANG-COG domain. Among children who have not had any chronic health issues, the proportion of children who were developmentally vulnerable was highest on the PHY-WELL and SOC domains and lowest on the LANG-COG domain.

Among children with any mental health issues, the highest proportion of children who were developmentally vulnerable was on the EMO and SOC domains, which aligns with previous research demonstrating strong associations with child's mental health issues.^{20,21} The lowest proportion of children who were developmentally vulnerable was on the LANG-COG domain. Among children who have not had any mental health issues, the proportion of children who were developmentally vulnerable was highest on the PHY-WELL and SOC domains and lowest on the LANG-COG domain.

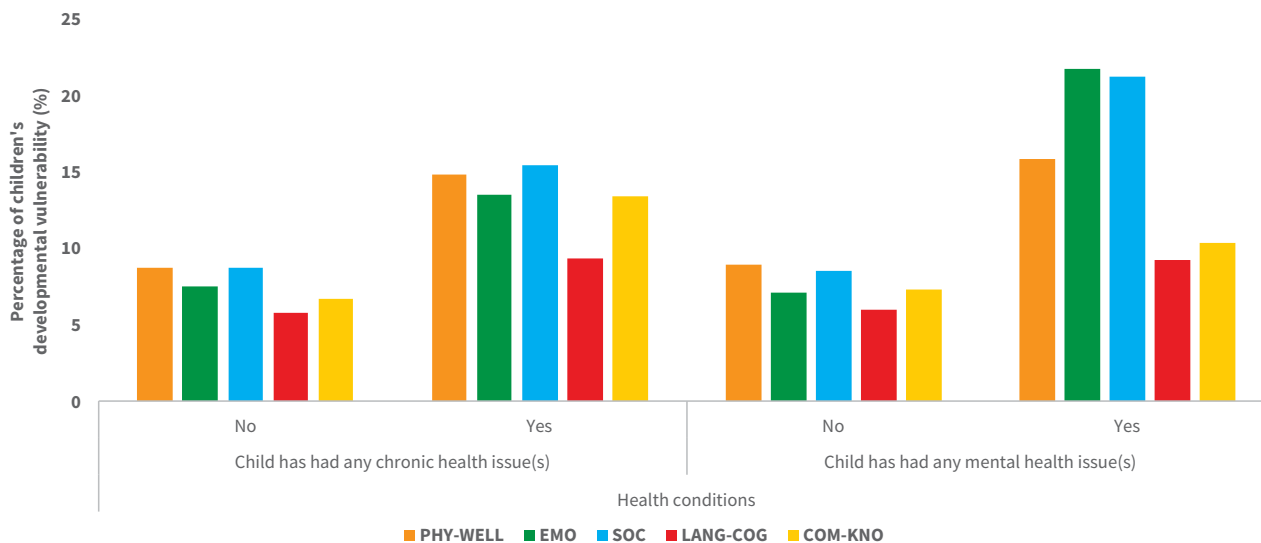


Figure 3. Percentage of children’s developmental vulnerability on specific domains by levels of the top two health condition indicators

Note. Top two indicators are based on the magnitude of association per disadvantage lens with children’s DV1 from our Phase One work. Abbreviations: COM-KNO communication skills and general knowledge; DV1 developmental vulnerability on one or more domain(s); EMO emotional maturity; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.

Geographic

The top two indicators within the geographic lens were ‘house overcrowding with one or more additional bedrooms needed’ and ‘rented tenure type’,²² which was based on the univariable associations with children’s DV1 in Phase One (Figure 4). Overall, the proportion of children who were developmentally vulnerable on each domain was higher among those experiencing house overcrowding and living in residences that were rented¹.

Among children who experienced house overcrowding, the highest proportion of children who were developmentally vulnerable was on the COM-KNO and PHY-WELL domains, followed by the SOC domain. The lowest proportion was on the EMO domain. Among children who did not experience house overcrowding, the proportion of children who were developmentally vulnerable was highest on the SOC domain and lowest on the LANG-COG domain.

Among children who lived in residences that were rented, the highest proportion of children who were developmentally vulnerable was on the PHY-WELL and SOC domains, followed by the COM-KNO and EMO domains. The lowest proportion was on the LANG-COG domain. Among children who lived in residences that were owned, the proportion of children who were developmentally vulnerable was highest on the COM-KNO domain and lowest on the LANG-COG domain.

1 The tenure category ‘living in residences that were rented’ also includes ‘occupied rent free’, ‘occupied under a life tenure scheme’, and ‘other tenure type’.

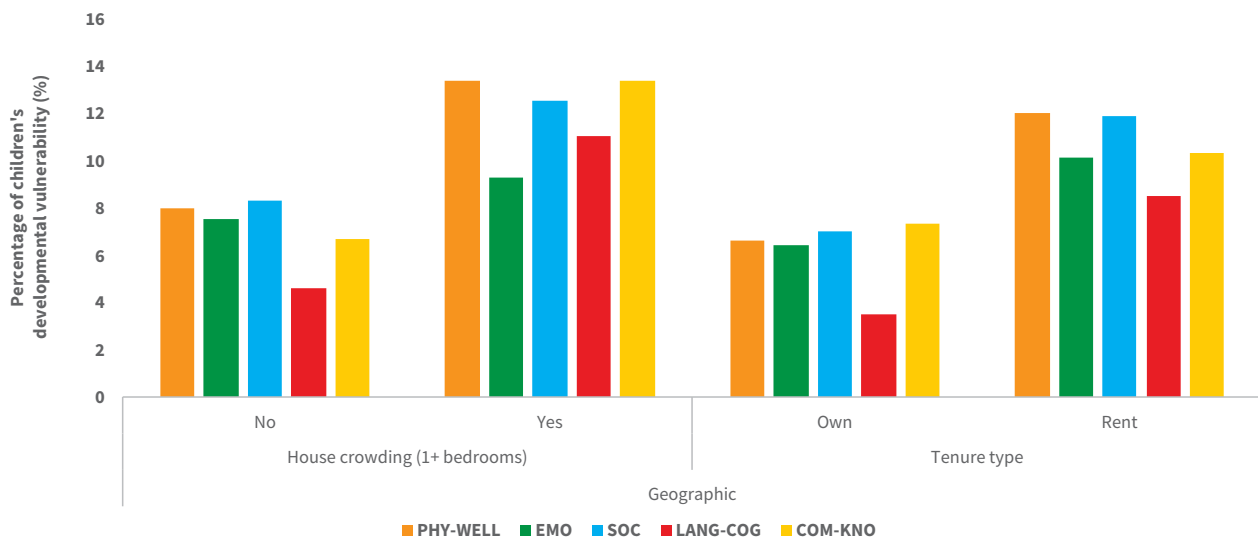


Figure 4. Percentage of children's developmental vulnerability on specific domains by levels of the top two geographic indicators

Note. Top two indicators are based on the magnitude of association per disadvantage lens with children's DV1 from our Phase One work. *Abbreviations:* COM-KNO communication skills and general knowledge; DV1 developmental vulnerability on one or more domain(s); EMO emotional maturity; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.

Risk factors

The top two indicators within the risk factors lens were 'child not regularly read to at home' and 'child born to a teenage mother', which was based on the univariable associations with children's DV1 in Phase One (Figure 5). Overall, the proportion of children who were developmentally vulnerable on each domain was higher among those who were not regularly read to at home and those who were born to a teenage mother.

Among children who were not regularly read to at home, the highest proportion of children who were developmentally vulnerable was on the PHY-WELL domain, followed by the LANG-COG, SOC, and COM-KNO domains. The lowest proportion of children developmentally vulnerable was on the EMO domain. Among children who were regularly read to at home, the proportion of children who were developmentally vulnerable was highest on the SOC domain and lowest on the LANG-COG domain (i.e., indicating that home reading was a protective factor for the LANG-COG domain).

Among children who were born to a teenage mother, the highest proportion of children who were developmentally vulnerable was on the SOC domain, followed by the PHY-WELL domain. The lowest proportion of children that were developmentally vulnerable was on the COM-KNO domain. Among children who were not born to a teenage mother, the proportion of children who were developmentally vulnerable was highest on the PHY-WELL and SOC domains and lowest on the LANG-COG domain.

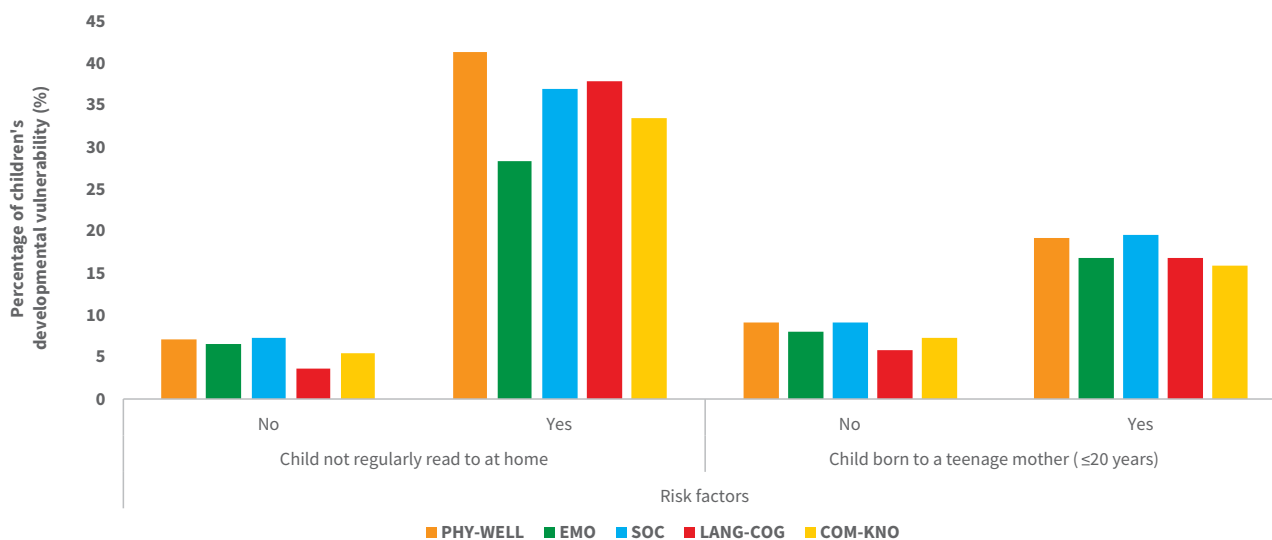


Figure 5. Percentage of children's developmental vulnerability on specific domains by levels of the top two risk factor indicators.

Note. Top two indicators are based on the magnitude of association per disadvantage lens with children's DV1 from our Phase One work. *Abbreviations:* COM-KNO communication skills and general knowledge; DV1 developmental vulnerability on one or more domain(s); EMO emotional maturity; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.

Priority populations

The top two indicators within the priority populations group were 'child not proficient in English' and 'parental ancestry' (Figures 6.1 and 6.2) as based on the magnitude of associations. Overall, the proportion of children who were developmentally vulnerable on each domain was higher among those who were not proficient in English (Figure 6.1) and for parental ancestry, the highest proportion was among those who had parents from Aboriginal and/or Torres Strait Islander backgrounds (Figure 6.2).

Among children who were not proficient in English, the highest proportion of children who were developmentally vulnerable was on the COM-KNO domain, which is consistent with previous research.^{23,24} The lowest proportion of children that were developmentally vulnerable was on the EMO domain. Among those who were proficient in English, the proportion of children who were developmentally vulnerable was highest on the SOC domain and lowest on the COM-KNO domain.

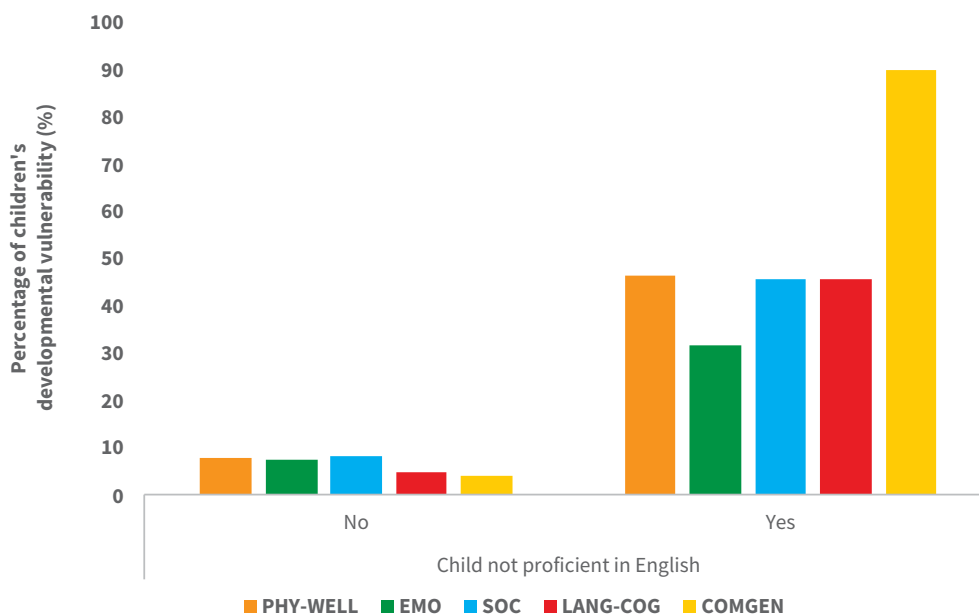


Figure 6. 1. Percentage of children’s developmental vulnerability on specific domains by levels of a top priority population indicator – Child not proficient in English

Note: Data on developmental vulnerability on COM-KNO were suppressed due to 90 per cent vulnerability rule as per the AEDC Data Guidelines section 7.2.2. *Abbreviations:* COM-KNO communication skills and general knowledge; EMO emotional maturity; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.

Among children who had parents that identify as from Anglo-Celtic backgrounds, the highest proportion of children who were developmentally vulnerable was on the PHY-WELL domain and lowest on the LANG-COG domain. Among children who have parents that identify as an Aboriginal and/or Torres Strait Islander, the highest proportion of children who were developmentally vulnerable was on the LANG-COG domain and lowest on the EMO domain. Among children who had parents that identify as South Central Asian, Northeast Asian, Southeast Asian backgrounds, this proportion of children who were developmentally vulnerable was highest on the COM-KNO domain and lowest on the LANG-COG domain.

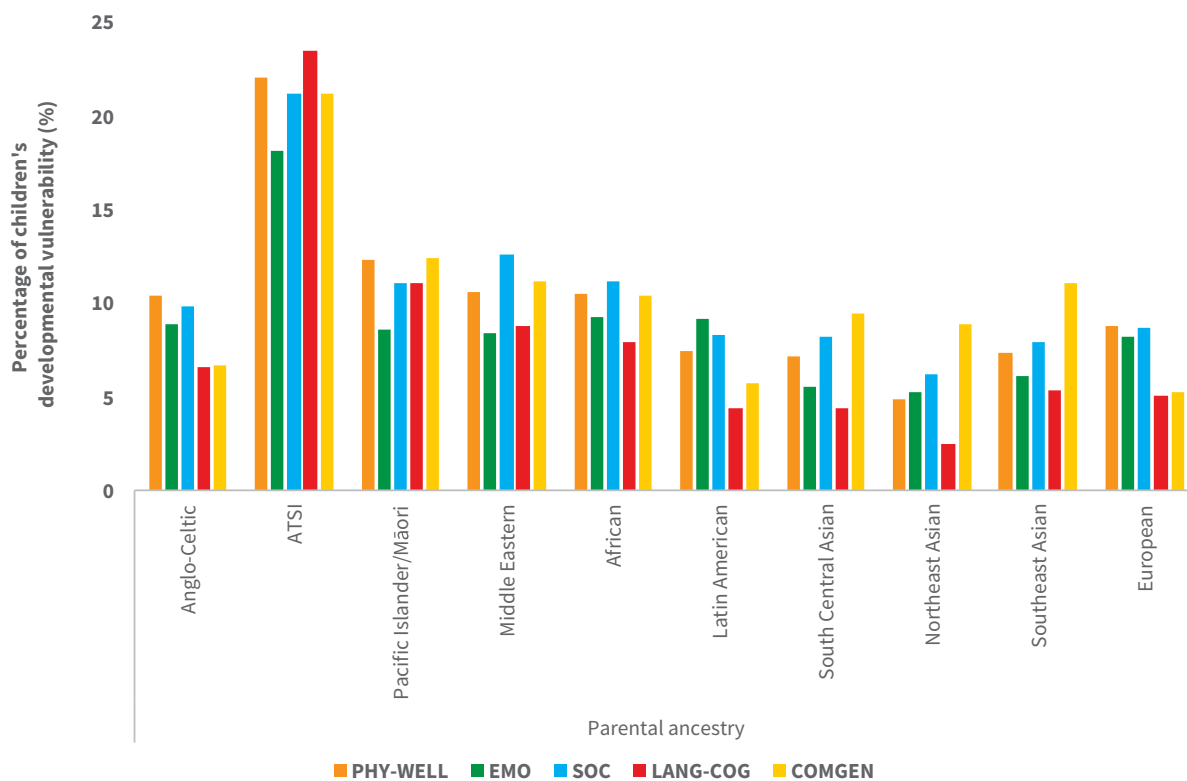


Figure 6.2. Percentage of children’s developmental vulnerability on specific domains by levels of a top priority population indicator – Parental ancestry

Abbreviations: COM-KNO communication skills and general knowledge; EMO emotional maturity; First Nations Aboriginal and/or Torres Strait Islander; LANG-COG language and cognitive skills (school-based); PHY-WELL physical health and wellbeing; SOC social competence.



Associations between child disadvantage indicators and children’s developmental vulnerability on specific domains

The univariable associations between each child disadvantage indicator and developmental vulnerability on each specific domain are shown in Appendix A, Table A.2. The 36 key child disadvantage indicators were then ranked according to the strength of univariable associations for each developmental vulnerability domain. The top 15 ranked child disadvantage indicators within each disadvantage lens for each domain are shown in Figures 7-11.

Figure 7 shows the top 15 ranked child disadvantage indicators across the four disadvantage lenses associated with the risk of children being developmentally vulnerable on the PHY-WELL domain. The indicator ‘child not regularly read to at home’ had the strongest univariable association with child PHY-WELL domain. Children who were not regularly read to at home had almost six times greater risk of being developmentally vulnerable on the PHY-WELL domain (RR=5.92, 95% CI: 5.79, 6.06) relative to children who were regularly read to at home. Sociodemographic indicators such as ‘Child Care Subsidy income thresholds’ and ‘parent highest education level’ were among the top 3 indicators associated with children’s developmental vulnerability on the PHY-WELL domain.

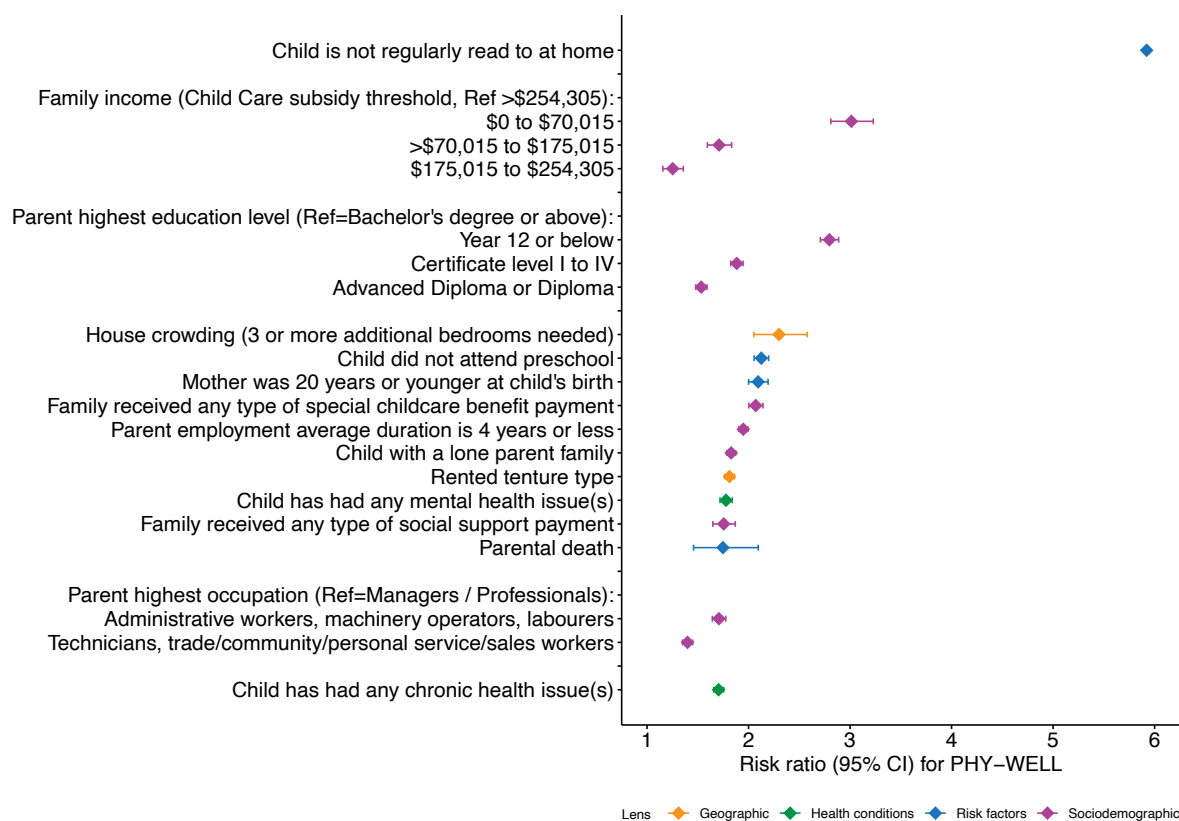


Figure 7. Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the PHY-WELL domain

Note. Based on magnitude of univariable associations. Abbreviation: CI Confidence Interval; PHY-WELL physical health and wellbeing.



Figure 8 shows the top 15 ranked child disadvantage indicators across the four disadvantage lenses associated with the risk of children being developmentally vulnerable on the EMO domain. The indicator ‘child not regularly read to at home’ had the strongest univariable association with the EMO domain. Children who were not regularly read to at home had slightly over four times greater risk of being developmentally vulnerable on the EMO domain (RR=4.29, 95% CI: 4.17, 4.42) relative to children who were regularly read to at home. Health condition indicator ‘child has had any mental health issues’ and sociodemographic indicator ‘Child Care Subsidy income thresholds’ were among the top 3 indicators associated with children’s developmental vulnerability on the EMO domain.

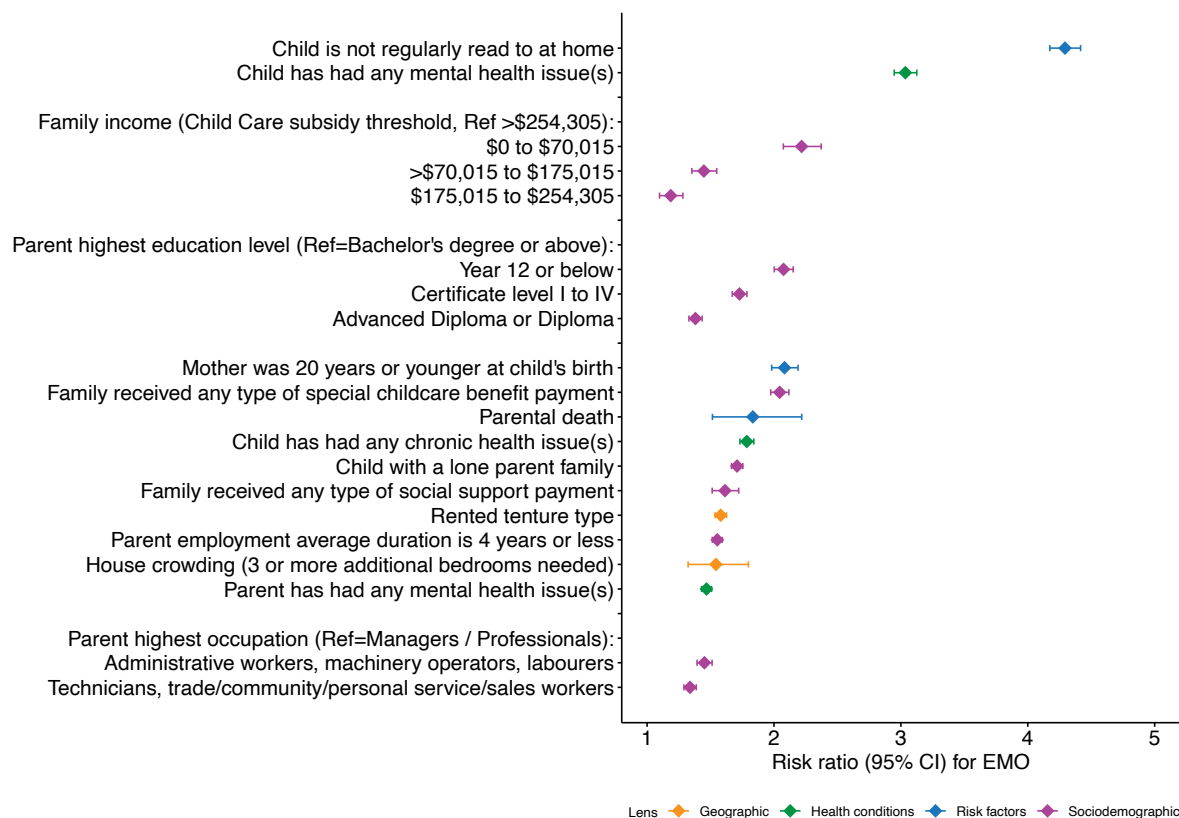


Figure 8. Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the EMO domain

Note. Based on magnitude of univariable associations. Abbreviation: CI Confidence Interval; EMO emotional maturity.



Figure 9 shows the top 15 ranked child disadvantage indicators associated with the risk of children being developmentally vulnerable on the SOC domain. The indicator ‘child not regularly read to at home’ had the strongest univariable association with the SOC domain. Children who were not regularly read to at home had almost five times greater risk of being developmentally vulnerable in the SOC domain (RR=5.12, 95% CI: 5.00, 5.24) relative to children who were regularly read to at home. Sociodemographic indicator ‘Child Care Subsidy income thresholds’ and health condition indicator ‘child has had any mental health issues for more than one year’ were among the top three indicators associated with children’s developmental vulnerability on the SOC domain.

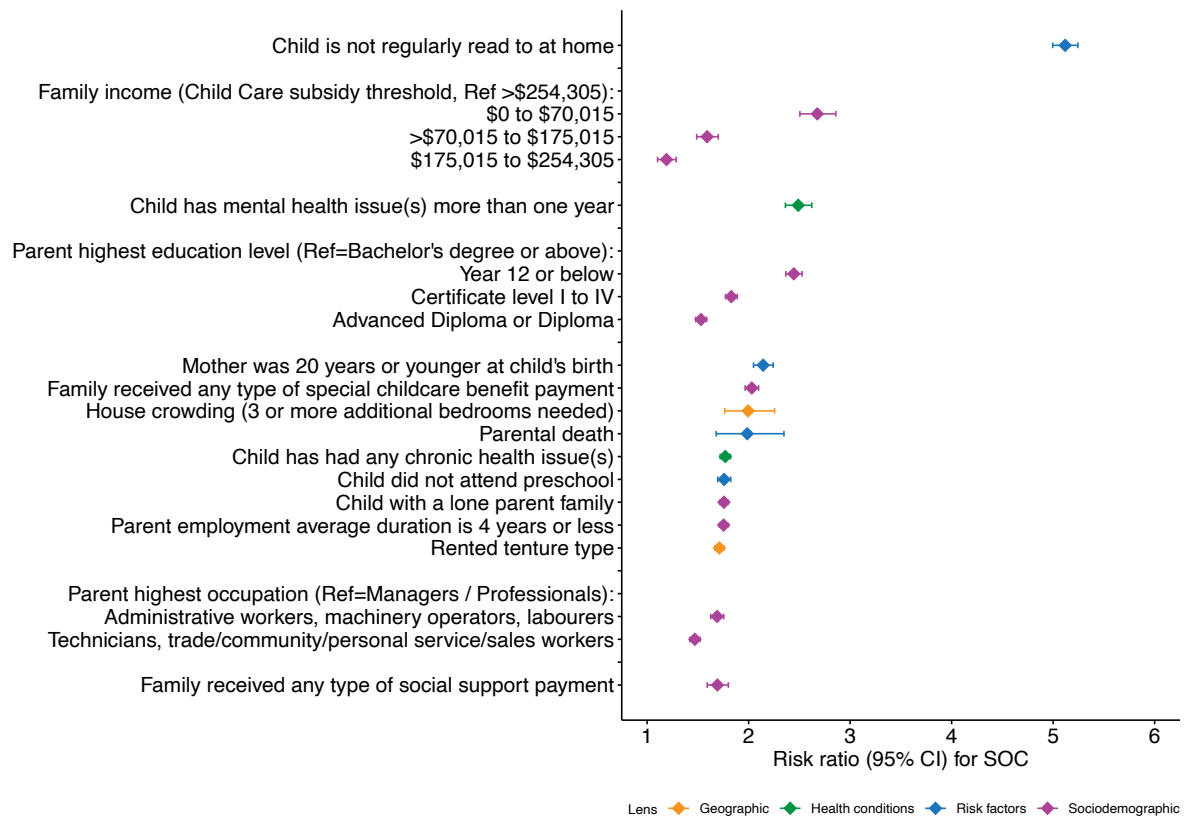


Figure 9. Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the SOC domain

Note. Based on magnitude of univariable associations. Abbreviation: CI Confidence Interval; SOC social competence.



Figure 10 shows the top 15 ranked child disadvantage indicators associated with the risk of children being developmentally vulnerable on the LANG-COG domain. The indicator ‘child not regularly read to at home’ had the strongest univariable association with the LANG-COG domain. Children who were not regularly read to at home had slightly over 10 times greater risk of being developmentally vulnerable in the LANG-COG domain (RR=10.16, 95% CI: 9.88, 10.44) relative to children who were regularly read to at home. Sociodemographic indicators ‘Child Care Subsidy income thresholds’ and ‘parent highest education level’ were among the top three indicators associated with children’s developmental vulnerability on the LANG-COG domain.

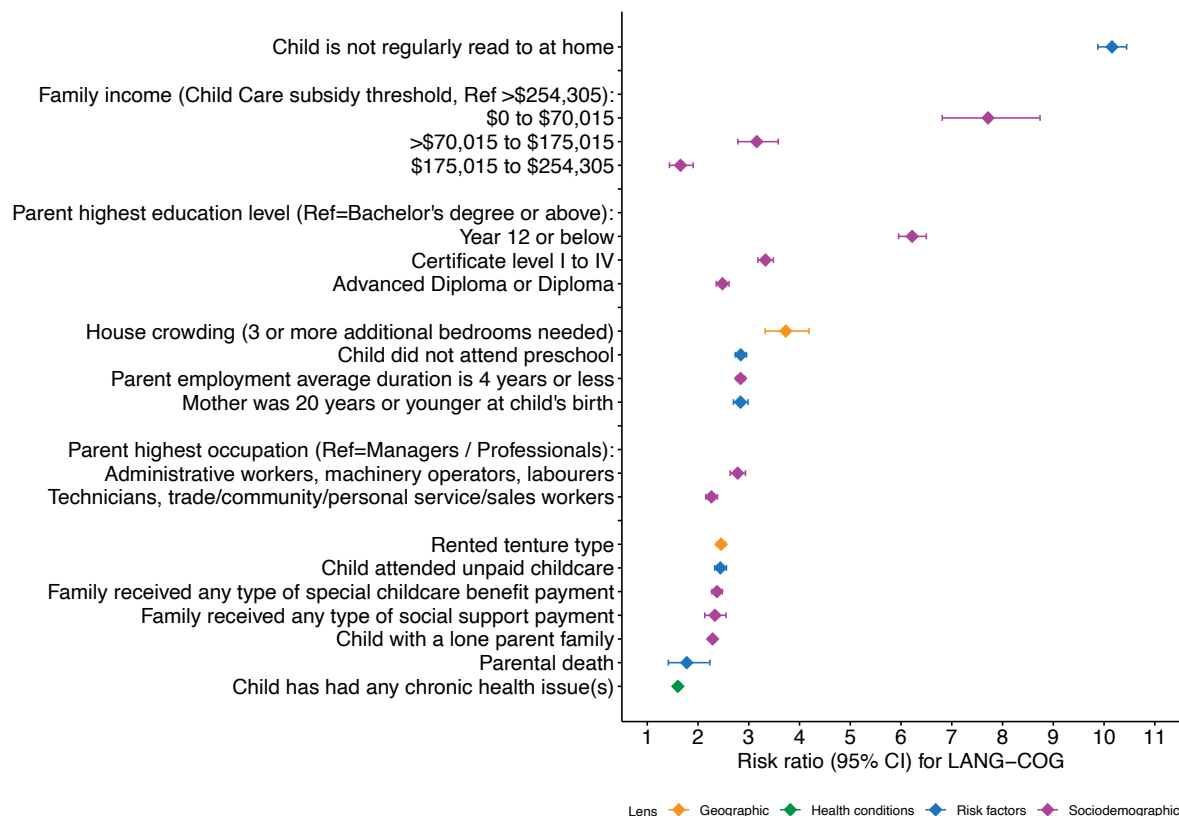


Figure 10. Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the LANG-COG domain

Note. Based on magnitude of univariable associations. Abbreviation: CI Confidence Interval; LANG-COG language and cognitive skills (school-based).



Figure 11 shows the top 15 ranked child disadvantage indicators associated with the risk of children being developmentally vulnerable on the COM-KNO domain. The indicator ‘child not regularly read to at home’ had the strongest univariable association with the COM-KNO domain. Children who were not regularly read to at home had slightly over 6 times greater risk of being developmentally vulnerable in the COM-KNO domain (RR=6.22, 95% CI: 6.06, 6.39) relative to children who were regularly read to at home. Sociodemographic indicators ‘Child Care Subsidy income thresholds’ and ‘parent highest education level’ were among the top 3 indicators associated with children’s developmental vulnerability on the COM-KNO domain.

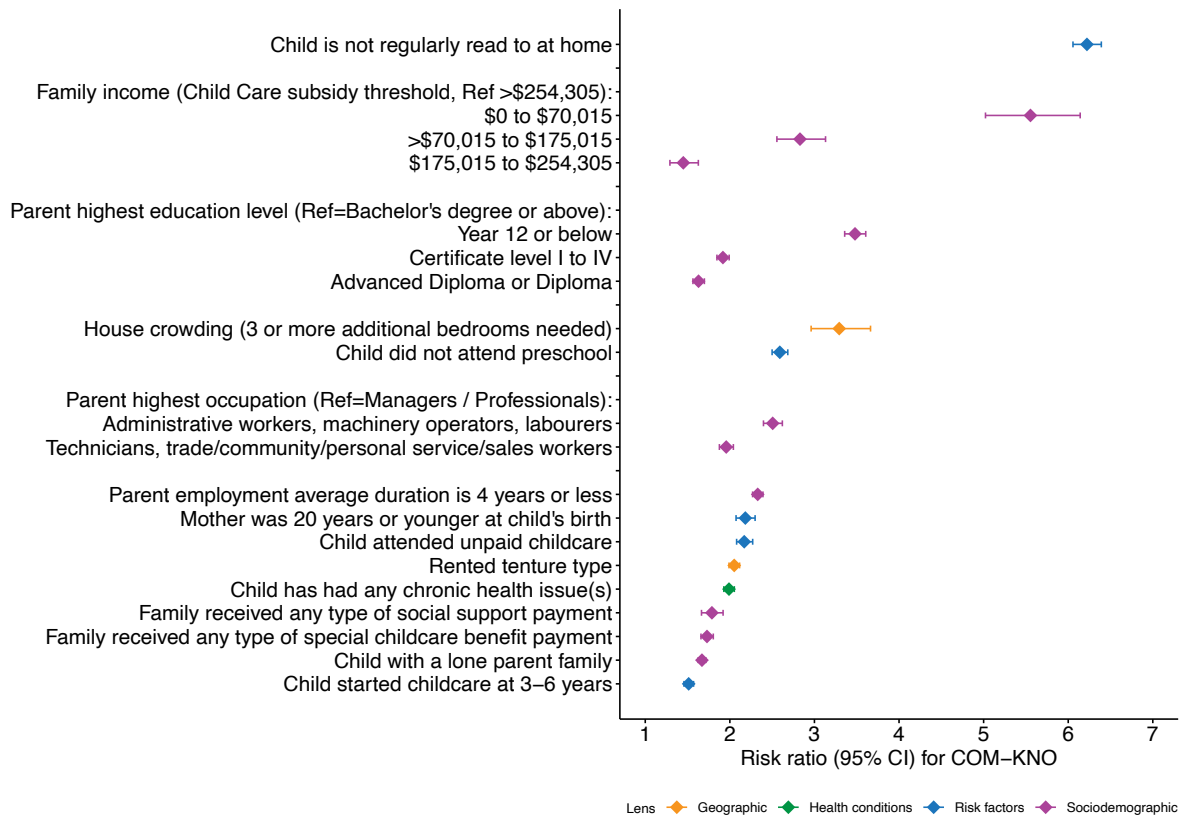


Figure 11. Top 15 ranked child disadvantage indicators associated with children’s developmental vulnerability on the COM-KNO domain

Note. Based on magnitude of univariable associations. Abbreviation: CI Confidence Interval; COM-KNO communication skills and general knowledge.



Associations between the priority population indicators and children's developmental vulnerability on specific domains

The univariable associations between each priority population indicator and each developmental vulnerability domain are shown in Table 3. The indicator 'child not proficient in English' had the strongest association with children's developmental vulnerability on each domain. Of note, the indicator 'child not proficient in English' was considered as a measured item contributing to one specific developmental domain (COM-KNO), so results should be interpreted with caution.

Table 3. Univariable associations between the priority population indicators and children's developmental vulnerability on specific developmental domains

Characteristics	PHY-WELL (N=25,816)	EMO (N=22,526)	SOC (N=25,970)	LANG-COG (N=17,065)	COM-KNO (N=20,625)
	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)
Child country of birth:					
Australia	Reference	Reference	Reference	Reference	Reference
Other English-Speaking country	0.98 (0.89, 1.08)	0.91 (0.81, 1.01)	0.96 (0.87, 1.06)	1.08 (0.96, 1.22)	1.06 (0.95, 1.18)
Other country	0.92 (0.86, 0.99)	0.96 (0.89, 1.03)	1.11 (1.05, 1.19)	1.14 (1.05, 1.23)	2.02 (1.91, 2.13)
Parent country of birth:					
Australia	Reference	Reference	Reference	Reference	Reference
Other English-Speaking country	0.77 (0.74, 0.80)	0.81 (0.78, 0.85)	0.80 (0.77, 0.83)	0.69 (0.65, 0.72)	0.91 (0.87, 0.95)
Other country	0.78 (0.75, 0.80)	0.75 (0.72, 0.77)	0.91 (0.89, 0.94)	0.86 (0.83, 0.89)	1.64 (1.59, 1.68)
Child Aboriginal and Torres Strait Islander status	2.34 (2.26, 2.42)	2.01 (1.93, 2.08)	2.16 (2.09, 2.23)	3.55 (3.43, 3.68)	2.54 (2.45, 2.64)
Parent ancestry with 10 categories:					
Anglo-Celtic ^a	Reference	Reference	Reference	Reference	Reference
Aboriginal and/or Torres Strait Islander	2.13 (1.96, 2.31)	2.04 (1.86, 2.23)	2.16 (1.98, 2.35)	3.56 (3.27, 3.87)	3.15 (2.89, 3.44)
Pacific Islander/Māori	1.19 (1.11, 1.28)	0.97 (0.89, 1.05)	1.14 (1.06, 1.23)	1.67 (1.55, 1.81)	1.85 (1.72, 2.00)
Middle Eastern	1.02 (0.96, 1.09)	0.94 (0.88, 1.01)	1.29 (1.21, 1.36)	1.33 (1.24, 1.43)	1.68 (1.57, 1.79)
African	1.01 (0.90, 1.13)	1.05 (0.93, 1.18)	1.15 (1.03, 1.28)	1.18 (1.03, 1.35)	1.55 (1.38, 1.74)
Latin American	0.71 (0.62, 0.83)	1.00 (0.87, 1.14)	0.85 (0.74, 0.98)	0.65 (0.54, 0.80)	0.86 (0.73, 1.02)
South Central Asian	0.69 (0.65, 0.73)	0.62 (0.58, 0.67)	0.83 (0.78, 0.88)	0.67 (0.62, 0.73)	1.41 (1.32, 1.50)
Northeast Asian	0.47 (0.44, 0.51)	0.58 (0.54, 0.63)	0.63 (0.59, 0.67)	0.38 (0.34, 0.42)	1.32 (1.24, 1.41)
Southeast Asian	0.70 (0.65, 0.75)	0.69 (0.64, 0.75)	0.81 (0.75, 0.87)	0.80 (0.73, 0.87)	1.66 (1.55, 1.77)
European	0.85 (0.82, 0.88)	0.92 (0.89, 0.96)	0.89 (0.86, 0.93)	0.77 (0.73, 0.80)	0.78 (0.75, 0.82)
Child LBOTE	0.96 (0.93, 0.99)	0.87 (0.84, 0.89)	1.12 (1.09, 1.15)	1.32 (1.28, 1.36)	2.23 (2.17, 2.29)
Parent LBOTE	0.88 (0.86, 0.91)	0.81 (0.78, 0.84)	1.05 (1.02, 1.08)	1.15 (1.10, 1.19)	2.04 (1.98, 2.10)
Child not proficient in English	5.85 (5.71, 5.99)	4.35 (4.21, 4.48)	5.72 (5.59, 5.86)	9.85 (9.59, 10.12)	22.76 (22.32, 23.2)

^a This group includes Anglo-Celtic and Australian backgrounds. Abbreviations: CI Confidence interval; COM-KNO Communication skills and general knowledge; EMO Emotional maturity; LANG-COG Language and cognitive skills (school-based); LBOTE Language background other than English; PHY-WELL Physical health and wellbeing; RR Risk ratio; SOC Social competence.



Summary of the top-ranked child disadvantage indicators and priority population indicators linked to specific developmental domains

All disadvantage indicators across the four disadvantage lenses and priority population groups identified in the top 15 shortlist for any one or more domains, are summarised in Table 4. There was a total of 18 indicators associated with one or more children's developmental vulnerability on specific domains; there were 7 indicators for the sociodemographic lens; 3 indicators for the health condition lens; 2 indicators for the geographic lens; 5 indicators for the risk factor lens, and one indicator for the priority population group.

Table 4. Summary of the child disadvantage and priority population indicators identified in the top 15 shortlist for developmental vulnerability on one or more specific domains

Sociodemographic (n=7)	Health conditions (n=3)	Geographic (n=2)	Risk factors (n=5)	Priority population (n=1)
<ul style="list-style-type: none"> Lower income household Lower parental highest education level Parent employment average duration was four years or less Child lived in a lone parent family Parent received any type of special childcare benefit Parent received any type of social support payment Parent highest occupation 	<ul style="list-style-type: none"> Child has had any chronic health issue(s) Child has had any mental health issue(s) Parent has had any mental health issue(s) 	<ul style="list-style-type: none"> House crowding (3 or more additional bedrooms needed) Rented tenure type 	<ul style="list-style-type: none"> Child is not regularly read to at home Mother's age at birth (teenager) Preschool non-attendance Parental death Child had no unpaid childcare 	<ul style="list-style-type: none"> Child not proficient in English



Conclusion

Key messages from Part One

- Univariable regression analyses were conducted to estimate the association of each child disadvantage and priority population indicator with children's developmental vulnerability on specific domains (PHY-WELL, EMO, SOC, LANG-COG and COM-KNO).
- The ranking of indicators by magnitude of univariable association with developmental vulnerability for each domain was broadly consistent with the ranking for DV1 in the Phase One report.
- Indicator 'child not regularly read to at home' had the strongest association with children's developmental vulnerability on all five domains.
- Indicator 'Child Care Subsidy income thresholds' was among the top three indicators on children's developmental vulnerability on all five domains, together with 'parental highest education level' on the PHY-WELL, LANG-COG and COM-KNO domains and 'child has had any mental health issues' on the EMO and SOC domains.

We found there was a clear pattern in the distribution of developmental vulnerability on each of the five specific domains by levels of the top two child disadvantage indicators within each disadvantage lens. Generally, children experiencing disadvantage had higher rates of developmental vulnerability across all five domains than their peers who did not experience disadvantage. Children who had parents that self-identified as an Aboriginal and/or Torres Strait Islander had higher rates of developmental vulnerability across all five domains than their peers who had parents from other backgrounds.

There was a total of 18 child disadvantage and priority population indicators that were associated with children's developmental vulnerability on the five specific domains, with the majority of factors within the sociodemographic lens. The ranking of disadvantage indicators by magnitude of the univariable associations with children's developmental vulnerability on each domain was broadly consistent with the ranking for DV1 in the Phase One report.¹¹

The top 15 indicators were almost the same 15 indicators across the five domains. In particular, 'child not regularly read to at home', 'Child Care Subsidy income thresholds', 'parent highest education level' were always in the top 4 indicators across the different disadvantage lens. There were also a few notable variations by domain, including: (1) child mental health issues were more strongly associated with developmental vulnerability on the EMO and SOC domains; (2) house crowding generally ranked highly but had a weaker association with developmental vulnerability on the EMO domain; (3) preschool non-attendance had a stronger association with developmental vulnerability on the PHY-WELL, LAN-COG and COM-KNO domains; and (4) employment duration, occupation and non-paid child care were more strongly associated with developmental vulnerability on the LAN-COG and COM-KNO domains.



Part 2

Comparison of child-level and area-level indicators of disadvantage

Overview

Area-level measures such as the Socio-Economic Indexes for Areas (SEIFA) are commonly used as a proxy indicator of an individual's socioeconomic status.²⁵ Compared to child-level measures of disadvantage (e.g., household income, parental education or occupation), SEIFA is a relative measure, which means that a given geographic area may be classified as disadvantaged relatively rather than disadvantaged in absolute terms.¹² The aim of Part 2 was to compare the child-level indicators of disadvantage considered in Phase One with the area-level SEIFA indicator and to estimate the univariable association of the area-level indicator with children's developmental vulnerability on one or more domain (DV1). We found that there were large differences between the child-level and area-level disadvantage indicators. We recommend careful consideration of the method of measuring disadvantage when investigating children's developmental vulnerability.

Methods

Measures

We focused on two types of disadvantage measures in this Part: 1) the 36 child-level indicators of disadvantage, which were identified in Phase One; and 2) an area-level indicator of disadvantage, SEIFA - Index of Relative Advantage and Disadvantage (IRSAD), which is widely used as a measure of relative socioeconomic position at the area level.²⁶ Based on the Australian Bureau of Statistics classification,²⁷ we categorised SEIFA into quintiles; with quintile 1 indicating the most disadvantaged and quintile 5 the least disadvantaged communities. We used developmental vulnerability on one or more domains (yes/no) (i.e., DV1), as the outcome of interest in this section. For further details, please refer to the [General Methods](#) section.

Statistical analysis

Cross-tabulations were used to summarise the distribution of each of the 36 child-level disadvantage indicators by area-level SEIFA quintiles.²⁷ This was conducted to explore differences in each of the child-level indicators among children living in communities with different socioeconomic levels. Consistent with the analysis performed in Phase One for the child-level disadvantage indicators, the univariable association of the area-level disadvantage measure with children's DV1 was estimated using univariable regression analyses with RRs (and corresponding 95% CIs) calculated relative to the least disadvantaged quintile (quintile 5) as the reference group.



Results

Child-level and area-level disadvantage indicators

A summary of the distribution of the child-level disadvantage indicators within the area-level SEIFA quintile groups are shown in Table 5. When using SEIFA as the area-level disadvantage indicator, almost one third (30.8%) of children were living in the most disadvantaged communities and 12.8% in the least disadvantaged communities. There was overlap between some of the child-level and area-level indicators, while others had very clear differences in distribution between the SEIFA quintiles.

In brief, there were higher proportions of children exposed to child-level disadvantage among those living in the most disadvantaged communities than those living in less disadvantaged communities. This suggests that children living in more disadvantaged communities are likely to be at higher risk of experiencing child-level disadvantage. For example, when considering eligibility for a Low Income Health Care Card as a child-level disadvantage indicator, 57.0% of children living in the most disadvantaged communities (quintile (Q) 1) came from families eligible for a Low Income Health Care Card, steadily decreasing to only 14.5% of children living in the least disadvantaged communities (Q5).

There were also considerable differences observed among children within the same SEIFA quintile. This suggests that using a single area-level disadvantage indicator like SEIFA may not accurately capture the complexity of child disadvantage. Children experiencing disadvantage were likely to be missed when using SEIFA as a single proxy measure. For example, of those children living in the most disadvantaged communities (Q1), 44.3% had family income in the lowest Family Tax Benefit A bracket of \$56,137 or less while 27.1% had family income in the highest bracket of \$99,864 or more. 27.1% of children were considered disadvantaged based on the area-level measure but least disadvantaged based on the child-level income measure. Similarly, of those children living in the least disadvantaged communities, 78.6% had family income in the highest bracket of \$99,864 or more while 10.8% had family income in the lowest bracket of \$56,137 or less. These 10.8% of children were considered least disadvantaged based on the area-level measure but disadvantaged on the child-level income measure.



Table 5. Cross tabulations between child-level and area-level disadvantage indicators

Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Total sample	270,748	12.8% (34,585)	16.0% (43,225)	18.6% (50,439)	21.8% (59,059)	30.8% (83,440)
Sociodemographic						
Poverty line:	262,398					
Above poverty line		30769 (92.3%)	37718 (89.9%)	43138 (88.0%)	49081 (85.3%)	66248 (81.7%)
Poverty line or below		2583 (7.7%)	4232 (10.1%)	5900 (12.0%)	8451 (14.7%)	14818 (18.3%)
Family eligible for a Low Income Health Care Card:	262,735					
No		28604 (85.5%)	33172 (79.1%)	35658 (72.7%)	36542 (63.6%)	34766 (43.0%)
Yes		4833 (14.5%)	8783 (20.9%)	13371 (27.3%)	20921 (36.4%)	46085 (57.0%)
Family Tax Benefit A, based on income group:	263,269					
Greater than \$99,864		26311 (78.6%)	28711 (68.3%)	29040 (59.1%)	27470 (47.7%)	21986 (27.1%)
\$56,137 to \$99,864		3526 (10.5%)	6918 (16.5%)	10315 (21.0%)	14636 (25.4%)	23236 (28.6%)
\$56,137 or less		3630 (10.8%)	6385 (15.2%)	9745 (19.8%)	15466 (26.9%)	35894 (44.3%)
Family Tax Benefit B, based on income group:	263,269					
Greater than \$100,900		26208 (78.3%)	28496 (67.8%)	28774 (58.6%)	27072 (47.0%)	21552 (26.6%)
\$100,900 or less		7259 (21.7%)	13518 (32.2%)	20326 (41.4%)	30500 (53.0%)	59564 (73.4%)
Child Care Subsidy income thresholds (4 categories):	263,273					
Greater than \$254,305		8664 (25.9%)	4432 (10.5%)	2396 (4.9%)	1403 (2.4%)	690 (0.9%)
Greater than \$175,015 to \$254,305		8307 (24.8%)	8681 (20.7%)	7448 (15.2%)	5540 (9.6%)	3207 (4.0%)
Greater than \$70,015 to \$175,015		11924 (35.6%)	20642 (49.1%)	26676 (54.3%)	30860 (53.6%)	32897 (40.6%)
\$0 to \$70,015		4572 (13.7%)	8259 (19.7%)	12580 (25.6%)	19769 (34.3%)	44322 (54.6%)
Child Care Subsidy income thresholds (2 categories):	263,269					
Greater than \$70,015		28895 (86.3%)	33755 (80.3%)	36520 (74.4%)	37803 (65.7%)	36794 (45.4%)
\$70,015 or less		4572 (13.7%)	8259 (19.7%)	12580 (25.6%)	19769 (34.3%)	44322 (54.6%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Parental highest education level:	257,257					
Bachelor's degree or above		25797 (78.0%)	26901 (64.6%)	25276 (52.2%)	22090 (39.2%)	17647 (22.7%)
Advanced Diploma or Diploma		3667 (11.1%)	6555 (15.8%)	9057 (18.7%)	11894 (21.1%)	16507 (21.2%)
Certificate level I to IV ^a		2426 (7.3%)	5655 (13.6%)	9801 (20.3%)	14705 (26.1%)	23791 (30.5%)
Year 12 or below		1167 (3.5%)	2502 (6.0%)	4249 (8.8%)	7605 (13.5%)	19965 (25.6%)
Parent highest occupation:	202,695					
Managers / Professionals		24086 (79.6%)	24910 (66.9%)	23297 (55.4%)	19927 (43.9%)	14594 (30.5%)
Technicians / Other types of workers ^b		3970 (13.1%)	8074 (21.7%)	11922 (28.4%)	15999 (35.3%)	19496 (40.8%)
Labourers / Others ^c		2213 (7.3%)	4239 (11.4%)	6804 (16.2%)	9425 (20.8%)	13739 (28.7%)
Parent highest occupation:	202,695					
White collar		27463 (90.7%)	30782 (82.7%)	31722 (75.5%)	30820 (68.0%)	28409 (59.4%)
Blue collar		2806 (9.3%)	6441 (17.3%)	10301 (24.5%)	14531 (32.0%)	19420 (40.6%)
Parent employment status:	213,737					
Employed		29911 (96.1%)	36524 (95.2%)	41191 (94.7%)	44268 (92.8%)	46302 (87.3%)
Not employed		1200 (3.9%)	1837 (4.8%)	2314 (5.3%)	3425 (7.2%)	6765 (12.7%)
Parent employment average duration:	268,420					
Greater than 4 years		28610 (83.9%)	34725 (81.2%)	38850 (77.8%)	42162 (72.0%)	44617 (53.8%)
4 years or less		5506 (16.1%)	8048 (18.8%)	11102 (22.2%)	16424 (28.0%)	38376 (46.2%)
Parent received social support payment:						
Age pension support payment:	270,753					
No		34575 (100.0%)	43202 (99.9%)	50383 (99.9%)	58979 (99.9%)	83131 (99.6%)
Yes		10 (0.0%)	23 (0.1%)	56 (0.1%)	80 (0.1%)	309 (0.4%)
Carer support payment:	270,748					
No		33490 (96.8%)	41355 (95.7%)	47831 (94.8%)	54955 (93.1%)	74027 (88.7%)
Yes		1095 (3.2%)	1870 (4.3%)	2608 (5.2%)	4104 (6.9%)	9413 (11.3%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Rent assistance support payment:	270,748					
No		31568 (91.3%)	35783 (82.8%)	36958 (73.3%)	35036 (59.3%)	30355 (36.4%)
Yes		3017 (8.7%)	7442 (17.2%)	13481 (26.7%)	24023 (40.7%)	53085 (63.6%)
Family support payment:	270,748					
No		6418 (18.6%)	4135 (9.6%)	3117 (6.2%)	2294 (3.9%)	1657 (2.0%)
Yes		28167 (81.4%)	39090 (90.4%)	47322 (93.8%)	56765 (96.1%)	81783 (98.0%)
Employment support payment:	270,748					
No		33860 (97.9%)	41366 (95.7%)	46819 (92.8%)	51529 (87.3%)	60744 (72.8%)
Yes		725 (2.1%)	1859 (4.3%)	3620 (7.2%)	7530 (12.7%)	22696 (27.2%)
Student support payment:	270,748					
No		34442 (99.6%)	42836 (99.1%)	49731 (98.6%)	57767 (97.8%)	80176 (96.1%)
Yes		143 (0.4%)	389 (0.9%)	708 (1.4%)	1292 (2.2%)	3264 (3.9%)
Disability support payment:	270,748					
No		34470 (99.7%)	42935 (99.3%)	49839 (98.8%)	57731 (97.8%)	79192 (94.9%)
Yes		115 (0.3%)	290 (0.7%)	600 (1.2%)	1328 (2.2%)	4248 (5.1%)
Any type of social security payments:	270,748					
No		6259 (18.1%)	4010 (9.3%)	3009 (6.0%)	2215 (3.8%)	1546 (1.9%)
Yes		28326 (81.9%)	39215 (90.7%)	47430 (94.0%)	56844 (96.2%)	81894 (98.1%)
Parent received special childcare benefit:						
At risk childcare benefit	203,773					
No		26508 (99.3%)	32672 (99.0%)	37463 (98.5%)	43439 (97.8%)	57968 (94.1%)
Yes		193 (0.7%)	326 (1.0%)	577 (1.5%)	996 (2.2%)	3631 (5.9%)
Financial hardship childcare benefit:	203,773					
No		26348 (98.7%)	32290 (97.9%)	36863 (96.9%)	42514 (95.7%)	57179 (92.8%)
Yes		353 (1.3%)	708 (2.1%)	1177 (3.1%)	1921 (4.3%)	4420 (7.2%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Grandparent childcare benefit:	203,777					
No		26695 (100.0%)	32982 (100.0%)	37983 (99.9%)	44295 (99.7%)	60697 (98.5%)
Yes		<10 (0.0%)	16 (0.0%)	57 (0.1%)	140 (0.3%)	902 (1.5%)
Jobs education and training childcare benefit:	203,773					
No		26583 (99.6%)	32614 (98.8%)	37273 (98.0%)	42513 (95.7%)	56091 (91.1%)
Yes		118 (0.4%)	384 (1.2%)	767 (2.0%)	1922 (4.3%)	5508 (8.9%)
Any special childcare benefit payments:	203,773					
No		26085 (97.7%)	31698 (96.1%)	35760 (94.0%)	40102 (90.2%)	49700 (80.7%)
Yes		616 (2.3%)	1300 (3.9%)	2280 (6.0%)	4333 (9.8%)	11899 (19.3%)
Child with a lone parent family:	266,614					
No		29527 (86.4%)	36062 (84.5%)	40522 (81.4%)	44006 (75.6%)	49444 (60.5%)
Yes		4636 (13.6%)	6637 (15.5%)	9265 (18.6%)	14206 (24.4%)	32309 (39.5%)
Household size with 6 or more people:	263,772					
5 people or less		32094 (95.7%)	40243 (95.6%)	46725 (94.9%)	54395 (94.3%)	76673 (94.4%)
6 people or more		1440 (4.3%)	1861 (4.4%)	2511 (5.1%)	3268 (5.7%)	4562 (5.6%)
Health conditions						
Parent has had any chronic health issue(s):	270,284					
No		25881 (74.9%)	30682 (71.1%)	34182 (67.9%)	38415 (65.2%)	51486 (61.8%)
Yes		8676 (25.1%)	12499 (28.9%)	16175 (32.1%)	20528 (34.8%)	31760 (38.2%)
Child has had any chronic health issue(s):	270,624					
No		30318 (87.7%)	37896 (87.7%)	44169 (87.6%)	51712 (87.6%)	73131 (87.7%)
Yes		4250 (12.3%)	5306 (12.3%)	6251 (12.4%)	7316 (12.4%)	10275 (12.3%)
Parent has had any mental health issue(s):	270,284					
No		15701 (45.4%)	19347 (44.8%)	22020 (43.7%)	23926 (40.6%)	30192 (36.3%)
Yes		18856 (54.6%)	23834 (55.2%)	28337 (56.3%)	35017 (59.4%)	53054 (63.7%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Parent mental health issue duration:	270,284					
Parent has no mental health issues or has had a mental health issue less than one year		22242 (64.4%)	27521 (63.7%)	31544 (62.6%)	34981 (59.3%)	44963 (54.0%)
Parent has had a mental health issue greater than one year		12315 (35.6%)	15660 (36.3%)	18813 (37.4%)	23962 (40.7%)	38283 (46.0%)
Child has had any mental health issue(s):	270,624					
No		32082 (92.8%)	40236 (93.1%)	46861 (92.9%)	54474 (92.3%)	76091 (91.2%)
Yes		2486 (7.2%)	2966 (6.9%)	3559 (7.1%)	4554 (7.7%)	7315 (8.8%)
Child mental health issue duration:	270,624					
Child has no mental health issues or has had a mental health issue less than one year		33970 (98.3%)	42539 (98.5%)	49593 (98.4%)	57961 (98.2%)	81625 (97.9%)
Child has had a mental health issue greater than one year		598 (1.7%)	663 (1.5%)	827 (1.6%)	1067 (1.8%)	1781 (2.1%)
Geographic						
House crowding (3 or more additional bedrooms needed):	222,297					
No		30222 (99.9%)	37580 (99.9%)	43001 (99.7%)	48579 (99.5%)	61691 (98.8%)
Yes		22 (0.1%)	56 (0.1%)	131 (0.3%)	262 (0.5%)	753 (1.2%)
House crowding (1 or more additional bedrooms needed):	222,297					
No		28852 (95.4%)	35725 (94.9%)	40242 (93.3%)	44453 (91.0%)	53169 (85.1%)
Yes		1392 (4.6%)	1911 (5.1%)	2890 (6.7%)	4388 (9.0%)	9275 (14.9%)
Dwelling type:	229,656					
Private dwellings		30852 (99.6%)	38411 (99.7%)	44218 (99.7%)	50349 (99.8%)	65176 (99.8%)
Collective dwellings		125 (0.4%)	122 (0.3%)	113 (0.3%)	123 (0.2%)	160 (0.2%)
Tenure type:	226,005					
Own		25311 (82.7%)	29616 (77.8%)	31461 (71.9%)	31161 (62.7%)	27803 (43.5%)
Rented ^d		5305 (17.3%)	8459 (22.2%)	12274 (28.1%)	18526 (37.3%)	36089 (56.5%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Child has moved residence address in the last 5 years:	223,469					
No		14093 (46.2%)	16556 (43.7%)	17276 (39.8%)	16909 (34.4%)	17050 (27.2%)
Yes		16395 (53.8%)	21287 (56.3%)	26134 (60.2%)	32233 (65.6%)	45536 (72.8%)
Risk factors						
Preschool non-attendance:	259,177					
No		32756 (96.6%)	40304 (95.8%)	46266 (94.7%)	52579 (93.0%)	69492 (89.3%)
Yes		1163 (3.4%)	1772 (4.2%)	2579 (5.3%)	3961 (7.0%)	8305 (10.7%)
Childcare non-attendance:	270,748					
No		26701 (77.2%)	32998 (76.3%)	38040 (75.4%)	44435 (75.2%)	61599 (73.8%)
Yes		7884 (22.8%)	10227 (23.7%)	12399 (24.6%)	14624 (24.8%)	21841 (26.2%)
Unpaid childcare:	243,738					
No		1082 (3.4%)	1505 (3.7%)	2058 (4.4%)	3098 (5.8%)	7472 (10.5%)
Yes		31060 (96.6%)	38647 (96.3%)	44600 (95.6%)	50564 (94.2%)	63652 (89.5%)
Child's age group at childcare entry:	203,773					
0-2 years		23388 (87.6%)	28224 (85.5%)	32005 (84.1%)	37139 (83.6%)	50289 (81.6%)
3-6 years		3313 (12.4%)	4774 (14.5%)	6035 (15.9%)	7296 (16.4%)	11310 (18.4%)
Child not regularly read to at home:	263,435					
No		33694 (98.7%)	41590 (97.9%)	47804 (96.8%)	54366 (94.5%)	69237 (86.7%)
Yes		445 (1.3%)	899 (2.1%)	1583 (3.2%)	3151 (5.5%)	10666 (13.3%)
Mother's age at birth (teenager):	267,145					
20 years and older		34117 (99.9%)	42542 (99.6%)	49382 (99.1%)	57002 (97.8%)	75807 (92.2%)
Younger than 20 years		39 (0.1%)	161 (0.4%)	439 (0.9%)	1282 (2.2%)	6374 (7.8%)
Mother's age at birth (later):	267,145					
Younger than 35 years		17018 (49.8%)	26413 (61.9%)	34752 (69.8%)	44339 (76.1%)	68115 (82.9%)
35 years or older		17138 (50.2%)	16290 (38.1%)	15069 (30.2%)	13945 (23.9%)	14066 (17.1%)



Child-level disadvantage indicator	N	Area-level SEIFA quintiles				
		Q5 (non-disadvantaged)	Q4	Q3	Q2	Q1 (most disadvantaged)
Parental death:	270,753					
No		34536 (99.9%)	43160 (99.8%)	50342 (99.8%)	58961 (99.8%)	83159 (99.7%)
Yes		49 (0.1%)	65 (0.2%)	97 (0.2%)	98 (0.2%)	281 (0.3%)

^a Certificate level I to IV including trade qualification. ^b Technicians, trade workers, community and personal service workers, and sales workers. ^c Clerical and administrative workers, machinery operators and driver, and labourers.

^d The tenure category “living in residences that were rented” also includes “occupied rent free”, “occupied under a life tenure scheme”, and “other tenure type”. * Abbreviations: SEIFA, socio-economic indexes for areas; Q quintiles.



Comparison of the magnitude of associations between child-level and area-level disadvantage indicators and DV1

Overall, there were higher proportions of children having developmental vulnerability on one or more domain(s) among those from more disadvantaged communities than those from least disadvantaged communities (see [Appendix B, Table B.1](#)). Children from the most disadvantaged communities had 2.4 times greater risk of being developmentally vulnerable on one or more domain(s) (RR=2.37, 95% CI: 2.30, 2.45) compared to their peers from the least disadvantaged communities (Figure 12, for values see [Appendix B, Table B.2](#)).

The estimated univariable associations of each of the child-level disadvantage indicators with DV1 (as reported in the Phase One report) are shown in [Appendix B, Table B.2](#) for comparison. A number of child-level indicators such as lower household income (<\$70,015, as compared to >\$254,305) and child not regularly read to at home had stronger univariable associations with DV1 than area-level SEIFA.

Our findings highlight that using SEIFA quintiles as a single measure of disadvantage is still useful to capture some aspects of disadvantage e.g., if services are directed at the community level. However, the issue with using area-level SEIFA is that it may not capture children who are disadvantaged at the child-level, so the distribution of the disadvantage group may be variable depending on the indicator. This indicates that children experiencing disadvantage who may need services may be missed.

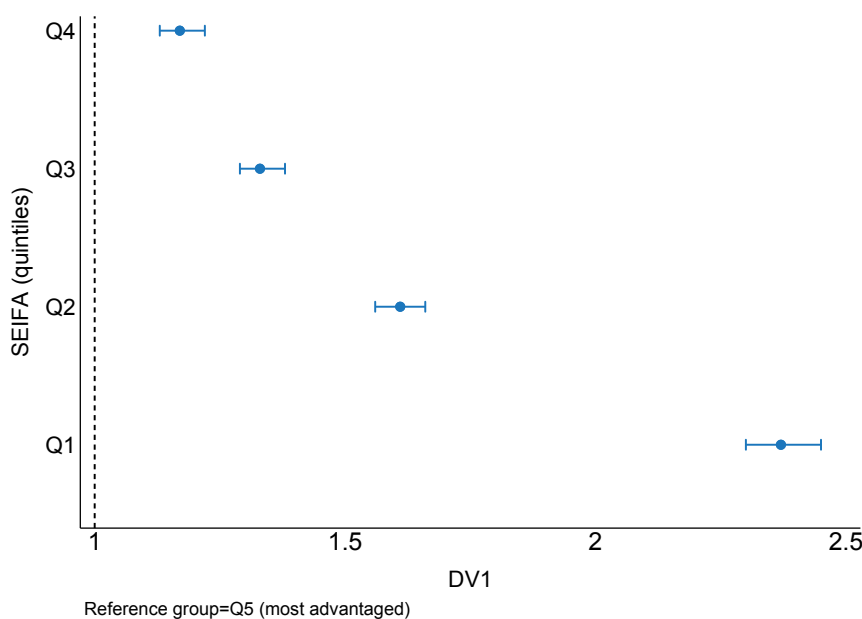


Figure 12. Association between SEIFA quintile groups and DV1

Note. Q1 indicates most disadvantage. *Abbreviations:* CI confidence interval; DV1 developmental vulnerability on one or more domain(s); SEIFA, socio-economic indexes for areas; RR risk ratio; Q quintiles.



Conclusion

Key messages from Part 2

- We observed clear differences in the distribution of each child-level disadvantage indicator across the SEIFA – IRSAD quintiles.
- There were higher proportions of children exposed to child-level disadvantage among those living in the most disadvantaged communities than those living in least disadvantaged communities, suggesting that children living in more disadvantaged communities are likely to be at higher risk of experiencing child-level disadvantage.
- However, there was considerable variation within SEIFA – IRSAD quintiles, therefore, area-level measures may not fully capture the distribution of disadvantage that children may experience at the individual level.

We found clear differences between children living in communities with different socioeconomic levels in terms of the distribution of child-level disadvantage indicators. We also found considerable variation in the child-level disadvantage indicators among children living in communities of the same socioeconomic level. This may suggest that area-based measures of disadvantage such as SEIFA–IRSAD, used as a proxy of child disadvantage, may not capture all aspects of individual disadvantage in children of their distribution.

Further, we also found that the univariable associations between child-level and area-level disadvantage indicators are largely variable in predicting children’s DV1. Child-level indicators such as lower Child Care Subsidy income thresholds and less home reading showed stronger associations with risk of DV1 than the SEIFA–IRSAD measure.



Part 3

Child-level disadvantage and priority population indicators for the prediction of children's developmental vulnerability

Overview

Advances in statistical modelling, such as machine learning, have shown good predictive performance in 'big data', which refer to the massive amounts and varieties of information in a structured or unstructured form.²⁸ Machine learning is an increasingly popular set of tools that allow computers to learn the relationship among numerical representations of data.²⁹ The purpose of Part 1 was to build upon the univariable analyses already completed in Phase One¹¹ and develop a multivariable prediction model comprised of the best combination of child-level disadvantage and priority population indicators for identifying groups of children at risk of DV1. Utilising the MADIP-FFY data, we reported on a subset of the 10 child-level disadvantage and priority population indicators that best predicted children's DV1.

Methods

Measures

The outcome of interest was developmental vulnerability on one or more domain(s) (i.e., DV1). The 36 child-level disadvantage indicators across the four disadvantage lenses and seven priority population indicators were considered as potential predictors.



Statistical analysis

Descriptive statistics and univariable associations of the child-level disadvantage and priority population indicators with the DV1 outcome were presented in the Phase One report.¹¹ Following on from our previous work, we provide an extension of the analysis of Phase One work over 3 steps: (1) predictor set selection, (2) identification of an optimal set of predictors; and (3) consideration of ensemble machine learning methods (Figure 13).

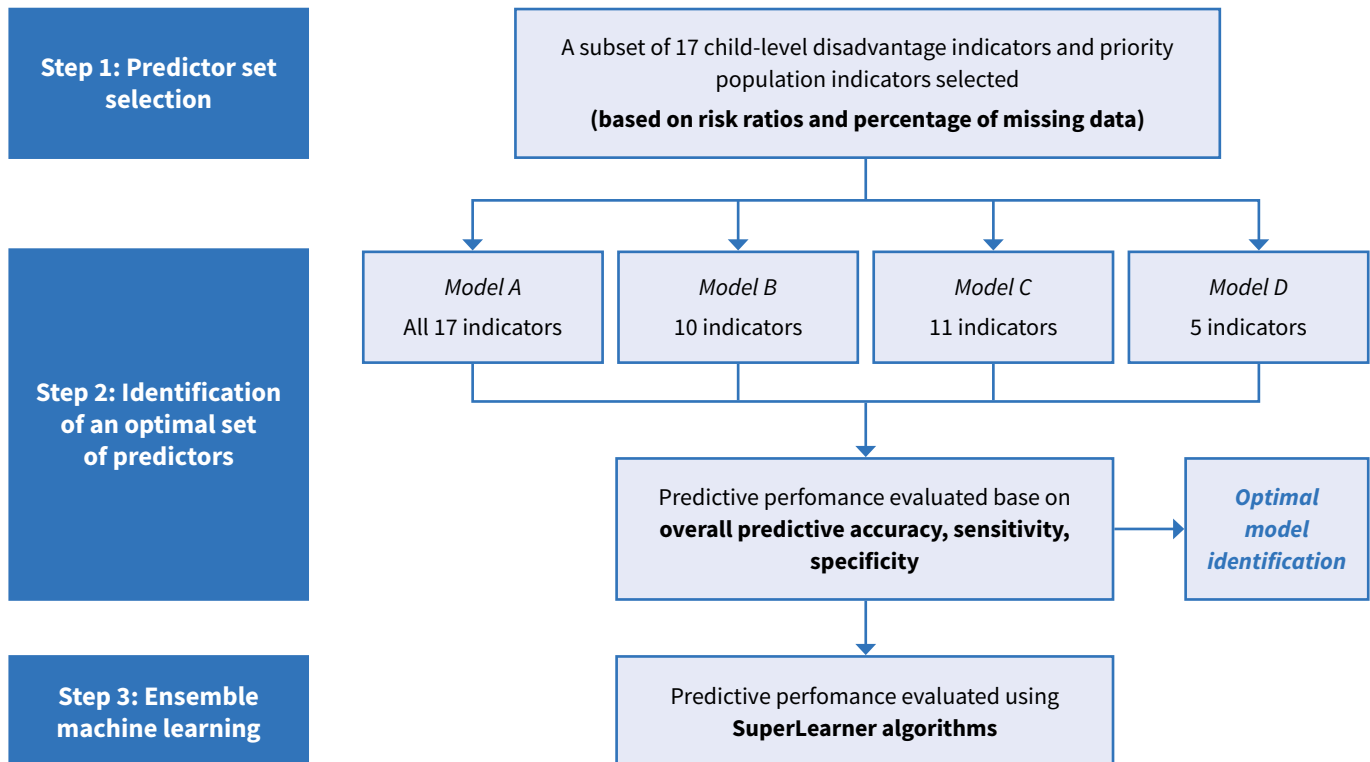


Figure 13. An overview of the methodology of Part 3 in this Phase Two report

Step 1: Predictor set selection

We used the findings from Phase One, specifically the magnitude of estimated univariable associations with DV1 (expressed as RR) and the percentage of missing observations to inform the selection of a subset of the 36 child-level disadvantage indicators and 7 priority population indicators for consideration in a multivariable prediction model.

Within each of the four disadvantage lenses, we ranked the relevant child-level disadvantage indicators in order of magnitude of the estimated univariable RR. Excluding any indicators with more than 30% missingness and based on our content expertise, we then selected between two and four of the highest ranked indicators within each lens, with the added constraint of selecting only one indicator within a given construct (e.g., income) to ensure adequate representation of all constructs within a given lens in the predictor set. ‘Constructs’ refers to the different groups of indicators within each lens. Each construct may be composed of multiple indicators (e.g., different ways of categorising household income i.e., Poverty line, Family Tax Benefit A and B, Child Care Subsidy income thresholds). These factors are strongly correlated by definitions, and selection of more than one variable within a construct would very likely yield similar results. We followed the same process to select three priority population indicators. Correlation plots describing associations between indicators within each disadvantage lens were also used to inform variable selection and presented for completeness.³⁰ However, multicollinearity is not a major concern when developing a multivariable prediction model given that it does not affect the statistics or predictions, but rather just the coefficients which were not interpreted in this case.³¹



Step 2: Identification of an optimal set of predictors

A series of candidate models were considered to identify the best combination of indicators for predicting the risk of children's developmental vulnerability.

These candidate models were chosen to reflect varying degrees of complexity while still ensuring adequate representation across the four disadvantage lenses and priority population groups. In summary:

- **Candidate model A** included all indicators identified in Step 1, i.e., four sociodemographic indicators, four health condition indicators, two geographic indicators, four risk factor indicators and three priority population indicators.
- **Candidate model B** included two child-level disadvantage indicators within each disadvantage lens and two priority population indicators with the strongest univariable associations with the DV1 outcome.
- **Candidate model C** included a total of 11 indicators, i.e., all the indicators within Candidate model B with an addition of the risk factor indicator 'preschool non-attendance' based on expert opinion and being one of the strongest child-level disadvantage univariable predictors from our Phase One work.
- **Candidate model D** included a total of five indicators i.e., the one child-level disadvantage indicator within each of the four disadvantage lenses and the one priority population indicator with the strongest univariable association with the DV1 outcome.

Logistic regression was used to fit candidate models to 60% of records selected randomly from the full analysis sample (referred to as the training data). Model predictive performance was assessed on the remaining 40% of records (referred to as the test data). This is a form of external model validation, aimed to avoid the undesirable property of overfitting, where the prediction model accurately reflects the specific features of the training data used to fit it but does not generalise well to new data. By splitting the analysis sample into the training and test datasets, we ensured that model predictive performance was objectively assessed on data that was not considered in building the model. This is a standard approach to building prediction models.

Model predictions representing the estimated probability of DV1 for each participant in the test data were obtained. Predicted outcomes (i.e., developmentally vulnerable on one or more domain(s) or not developmentally vulnerable on one or more domain(s)) were then derived from these estimated probabilities with reference to a probability threshold of 0.5.³² For example, all records with estimated probabilities less than the probability threshold were classified as having a predicted outcome of 'not developmentally vulnerable', while records with estimated probabilities of greater than or equal to the probability threshold were classified as having a predicted outcome of 'developmentally vulnerable'. These predicted developmental vulnerability outcomes were then compared to the actual observed developmental vulnerability outcomes to assess model predictive performance within the test dataset.

The following measures of predictive performance were considered:

- *Overall predictive accuracy*: the proportion of records in the test dataset where the value of the outcome predicted from the model matched the observed value of the outcome. In other words, the proportion of children for whom the model correctly classified developmental vulnerability status (as either developmentally vulnerable or not developmentally vulnerable).
- *Sensitivity*: the proportion of children in the test dataset with a known or observed outcome status of developmentally vulnerable who were correctly identified (or predicted) as being developmentally vulnerable by the model.
- *Specificity*: the proportion of children in the test dataset with a known or observed outcome status of not developmentally vulnerable who were correctly identified (or predicted) as being not developmentally vulnerable by the model.

These measures of predictive performance are specific to a particular value of the probability threshold described above. The area under the receiver operating characteristic curve (AUC) is another measure of model overall predictive performance that summarises sensitivity and specificity at all possible values of the probability threshold. AUC values range between 0 and 1 and values closest to 1 are indicative of better performance. The AUC also enables determination of the optimal probability threshold that maximises sensitivity and specificity.



For each of the candidate models, the overall predictive performance measure of AUC and measures of sensitivity, specificity and overall predictive accuracy calculated at a probability threshold of 0.5 were obtained with 95% CI. The parsimonious model was chosen as the one with the *greatest predictive performance* based on these measures and the number of predictors used within a model.

Step 3: Consideration of ensemble machine learning methods

We used the SuperLearner prediction algorithm,³³ available in the statistical program R, to explore the potential gains in predictive accuracy that could be achieved using a more complex machine learning approach applied to the optimal set of indicators identified in Step 2. SuperLearner is a flexible machine-learning algorithm that offers a wide range of individual algorithms or 'base learners', for developing prediction models.³³ We specified three different popular learners: (i) *SL.glm* refers to fitting a general linear model and is analogous to a standard multivariable regression model as performed in Step 2; (ii) *SL.randomForest*, a well-known machine learning approach for prediction that utilises decision trees; and (iii) *SL.glmnet* refers to elastic net regression for prediction combined with LASSO and ridge regressions.

Each of the individual prediction algorithms were evaluated in terms of the same predictive performance measures described in Step 2 and then weighted accordingly to produce a new "ensemble" prediction algorithm expected to perform at least as well as any of the individual methods.³³ Predictive performance measures were estimated within SuperLearner using 10-fold cross-validation. This is an extension of the (2-fold) external model validation approach described in Step 2, which this time, splits the data into ten blocks, applies the model-fitting algorithm on nine of the ten blocks (the training data) and then obtains predictions for the remaining one block of test data. This process is repeated until all ten blocks have served as test data. Estimates of overall predictive performance were compared to those obtained in Step 2.

Results

Key findings

- The best combination of child-level disadvantage indicators for predicting children's developmental vulnerability was identified.
- The chosen combination of indicators included:
 - Two **sociodemographic** indicators: household income (using the childcare subsidy income thresholds) and parental highest education level
 - Two **health conditions** indicators: child's chronic health issues and child's mental health issues
 - Two **geographic** indicators: house overcrowding and rented tenure type
 - Two **risk factors** indicators: child not regularly read to at home and child born to a teenage mother
 - Two **priority population** indicators: child not proficient in English and parental ancestry.
- However, a prediction model based on these disadvantage indicators alone is not sufficient to accurately quantify the risk of children's developmental vulnerability.
- Prediction of children's developmental vulnerability warrants the use of a diverse range of disadvantage indicators from different lenses, given the complex multi-faceted nature of child disadvantage and the consideration of other important factors outside of the measured aspects of disadvantage that are important in determining children's developmental vulnerability.



Associations between the child disadvantage and priority population indicators and DV1

Based on the estimated univariable associations with DV1 and the proportion of missingness, 17 indicators were selected for consideration in a multivariable prediction model (Table 6). For a full list of all estimated univariable associations with developmental vulnerability for each of the 36 key disadvantage indicators and priority population indicators, see [Appendix C, Tables C.1-5](#). Correlation plots confirmed expected associations between child-level disadvantage and priority population indicators within each disadvantage lens ([Appendix C, Figures C.1-5](#)).

Table 6. Summary of the 17 candidate child disadvantage and priority population indicators as predictors of DV1 ordered by magnitude of univariable association

Indicators	Categories	DV1	% Missing of total N=293,910
		RR (95% CI)	
Sociodemographic			
Child Care Subsidy income thresholds	0: \$254,305 or more;	Reference	10.3%
	1: \$175,015 to \$254,305;	1.19 (1.13, 1.25)	
	2: \$70,015 to \$175,015;	1.62 (1.55, 1.69)	
	3: \$70,015 or less	2.60 (2.49, 2.71)	
Parental highest education level	0: Bachelor's degree or above	Reference	11.6%
	1: Advanced Diploma or Diploma	1.47 (1.44, 1.50)	
	2: Certificate level I to IV (including trade qualification);	1.72 (1.68, 1.75)	
	3: Year 12 or below	2.37 (2.33, 2.42)	
Child with a lone parent family	0: No;	Reference	9.2%
	1: Yes	1.61 (1.59, 1.64)	
Parent received social support payment	0: No;	Reference	6.9%
	1: Yes	1.57 (1.51, 1.63)	
Health conditions			
Child has had any mental health issue(s)	0: No	Reference	7.0%
	1: Yes	1.82 (1.78, 1.86)	
Child has had any chronic health issue(s)	0: No	Reference	7.0%
	1: Yes	1.59 (1.56, 1.62)	
Parent mental health issue duration	0: Parent has no mental health issues or has had a mental health issue less than one year;	Reference	7.1%
	1: Parent has had a mental health issue greater than one year	1.24 (1.23, 1.26)	
Parent has had any chronic health issue(s)	0: No	Reference	7.1%
	1: Yes	1.23 (1.21, 1.25)	



Indicators	Categories	DV1	% Missing of total N=293,910
		RR (95% CI)	
Geographic			
Tenure type	0: Own (Owned outright, Owned with a mortgage, Being purchased under a shared equity scheme);	Reference	22.3%
	1: Rented (Rented, Being occupied rent-free, Being occupied under a life tenure scheme, Other tenure type)	1.64 (1.62, 1.67)	
House crowding (1 or more additional bedrooms needed)	0: None needed/Spare bedrooms	Reference	23.6%
	1: One or more extra bedrooms needed	1.58 (1.55, 1.62)	
Risk factors			
Child is regularly read to at home	1: Yes (Very true or somewhat true)	Reference	9.5%
	0: No (Not Sure)	4.06 (4.00, 4.11)	
Mother's age at birth (teenager)	0: >=20 years	Reference	8.2%
	1: <20 years	1.92 (1.87, 1.97)	
Preschool attendance	1: Yes	Reference	10.9%
	0: No	1.79 (1.75, 1.82)	
Parental death	0: No	Reference	6.9%
	1: Yes	1.59 (1.42, 1.78)	
Priority populations			
Child not proficient in English	0: No (Very good, good, average)	Reference	0.1%
	1: Yes (Very poor or poor)	5.29 (5.24, 5.34)	
Parent ancestry	0: Anglo-Celtic ^a	Reference	16.8%
	1: Aboriginal and/or Torres Strait Islander	2.03 (1.93, 2.13)	
	2: Pacific Islander/Māori	1.26 (1.20, 1.32)	
	3: Middle Eastern	1.19 (1.15, 1.24)	
	4: African	1.15 (1.08, 1.23)	
	5: Latin American	0.86 (0.79, 0.94)	
	6: South Central Asian	0.87 (0.84, 0.90)	
	7: Northeast Asian	0.76 (0.73, 0.79)	
	8: Southeast Asian	0.94 (0.90, 0.98)	
9: European	0.88 (0.86, 0.90)		
Parents LBOTE	0: No	Reference	15.8%
	1: Yes	1.13 (1.11, 1.15)	

^a This group includes Anglo-Celtic and Australian backgrounds. Abbreviations: CI Confidence Interval; DV1 developmental vulnerability on one or more domain(s); LBOTE Language background other than English; RR Risk ratio.



Identification of optimal set of predictors

The child-level disadvantage and priority population indicators included in each of the four candidate models are shown in Table 7.

Table 7. Summary of the child disadvantage and priority population indicators in the four candidate models for the prediction of DV1

Groups	Indicators	Model A	Model B	Model C	Model D
Sociodemographic	Child Care Subsidy income thresholds	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Parental highest education level	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Parent received social support payment	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Child with a lone parent family	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Health conditions	Child has had any chronic health issue(s)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Parent has had any chronic health issue(s)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Child has had any mental health issue(s)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Parent mental health issue duration	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Geographic	House crowding (1 or more additional bedrooms)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Rented tenure type	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Risk factors	Preschool non-attendance	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Child is not regularly read to at home	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Mother's age at birth (teenager)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Parental death	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Priority populations	Parents LBOTE	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Child not proficient in English	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Parental ancestry	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Total indicators	17	10	11	5

Abbreviation: LBOTE Language background other than English.

The predictive performance of the four candidate models based on the AUC, sensitivity, specificity, and overall predictive accuracy are summarised in Table 8.



Table 8. Predictive performance of the four candidate models using 60% training and 40% test data (2-fold cross validation)

Candidate model	AUC (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Predictive Accuracy (95% CI)
Model A: 17 indicators	0.72(0.71, 0.72)	0.98 (0.98, 0.99)	0.26 (0.26, 0.27)	0.86 (0.85, 0.86)
Model B: 10 indicators	0.72 (0.72, 0.72)	0.98 (0.98, 0.98)	0.27 (0.27, 0.28)	0.85 (0.85, 0.86)
Model C: 11 indicators	0.72 (0.72, 0.72)	0.98 (0.98, 0.98)	0.27 (0.26, 0.27)	0.86 (0.85, 0.86)
Model D: 5 indicators	0.66 (0.65, 0.66)	0.98 (0.98, 0.98)	0.17 (0.17, 0.17)	0.83 (0.83, 0.83)

Abbreviation: AUC area under the curve, CI confidence intervals.

Candidate Model B, with a total of 10 indicators, included the sociodemographic indicators of household income (using the Child Care Subsidy income thresholds) and highest education level of parents, the health condition indicators include child's chronic health issues and child's mental health issues, the geographic indicators include house overcrowding and living in residences that were rented, the risk factors indicators include child not regularly read to at home and child born to a teenage mother, and the priority population indicators include child not proficient in English and parental ancestry.

Model B achieved the parsimonious model with strong overall performance for predicting the risk of children's DV1, given that it has lowest numbers of predictors compared to Model A and C (17 and 11 indicators, respectively). The overall predictive accuracy of Model B was 85%, indicating that developmental vulnerability status was correctly classified for 85% of children in the test dataset. This was largely driven by a very high specificity of 98%, indicating that the model correctly identified 98% of children who were known or observed to be not developmentally vulnerable. Conversely, the model was much less successful at accurately identifying children who were developmentally vulnerable, doing so for only 27% of those children known or observed to be developmentally vulnerable. The remaining 73% of children observed as developmentally vulnerable were incorrectly classified by the prediction model not developmentally vulnerable.

While the 10 indicators in candidate Model B were considered the best combination of indicators for predicting DV1, this model did not perform well in quantifying the risk of DV1 in children in their first year of school. While the addition of the preschool attendance indicator in Model C produced similar predictive performance, Model B was considered optimal given its slightly higher sensitivity.

Consideration of ensemble machine learning methods

A comparison of the predictive performance of the SuperLearner individual learners and the ensemble algorithm for Candidate model B are shown in Table 9. The results showed that the machine learning algorithms did not achieve substantial gains in predictive performance relative to multivariable logistic regression, with the predictive performance measures of the ensemble algorithms comparable to that of the *glm* (logistic regression) approach.



Table 9. Predictive performance of SuperLearner individual learners and ensemble algorithm for the best candidate model

Candidate model	Algorithm	AUC (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Prediction Accuracy (95% CI)
Model B: 10 indicators	SL (ensemble)	0.721 (0.719, 0.722)	0.982 (0.981, 0.983)	0.269 (0.265, 0.271)	0.852 (0.851, 0.854)
	SL.glm	0.721 (0.651, 0.722)	0.982 (0.981, 0.983)	0.269 (0.265, 0.273)	0.853 (0.851, 0.854)
	SL.glmnet	0.721 (0.654, 0.723)	0.982 (0.981, 0.983)	0.269 (0.852, 0.854)	0.853 (0.852, 0.854)
	SL.randomForest	0.653 (0.603, 0.722)	0.982 (0.983, 0.992)	0.269 (0.243, 0.273)	0.852 (0.817, 0.854)

Abbreviation: AUC area under the curve; CI confidence interval; SL SuperLearner.

Conclusion

Key messages from Part 3

- A subset of the 36 child-level disadvantage indicators and 7 priority population indicators identified from Phase One was selected based on strength of estimated univariable associations with DV1 and the percentage of missing observations.
- Four candidate prediction models were established, each with a different selection of indicators across the four disadvantage lenses and priority population groups, to identify the best combination of indicators for predicting the risk of children's DV1.
- Model B with 10 child disadvantage and priority population indicators achieved the optimal predictive performance based on the AUC, sensitivity, specificity, and overall predictive accuracy.

The parsimonious model for the prediction of DV1 was Model B with 10 child-level disadvantage and priority population indicators (Figure 14).

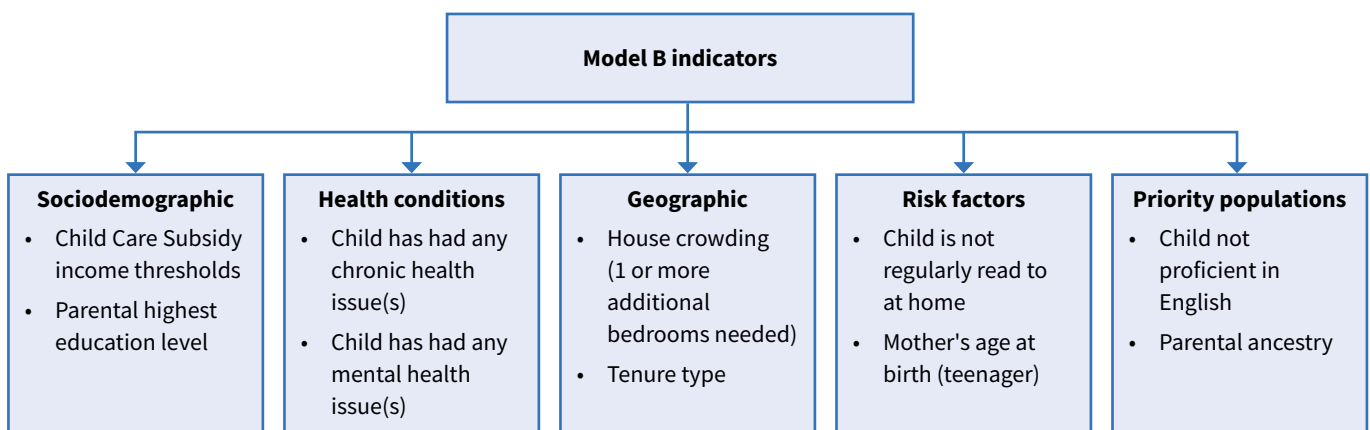


Figure 14. The 10 child-level disadvantage and priority population indicators in candidate Model B



The model's overall predictive accuracy of 85% was high indicating generally good performance, however this was largely driven by the model's ability to accurately identify children who were not developmentally vulnerable with a specificity of 98%. These children comprised approximately 80% of the dataset. Of the minority of children who were developmentally vulnerable, the model was only able to identify these children correctly 27% of the time, while the remaining 73% were incorrectly classified as not developmentally vulnerable. This pattern of high specificity and low sensitivity is not uncommon when the data are unbalanced, that is, when only a small proportion of the sample are observed to have the outcome of interest, in this case, 20% of children observed as developmentally vulnerable. Together with the performance measures of the optimal model, findings show that while representing the optimal set of predictors, this collection of indicators alone was not sufficient to accurately and reliably distinguish the small proportion of children who were developmentally vulnerable from the majority who were not. As a result, the overwhelming majority of children were predicted or classified into the most prevalent category, which in this case was not developmentally vulnerable (thus leading to the high specificity and low sensitivity).

Although the 10 indicators identified in the optimal set represent aspects of disadvantage most likely to identify children at risk of developmental vulnerability, a prediction model based on these disadvantage indicators alone was not sufficient to accurately estimate the risk of developmental vulnerability in children in their first year of school. This suggests that there are other important determinants of developmental vulnerability in children at school entry that were not captured in the prediction model including both unmeasured aspects of disadvantage and factors beyond disadvantage, such as genetic factors, prenatal biological factors and children's inherent intellectual and motivation levels.

Our results showed that the use of complex machine learning algorithms did not improve the performance of the optimal set of indicators in predicting DV1. This can most likely be explained by the fact that these predictors are almost all binary indicator variables and as such, there are no decisions to be made about how to include these variables in a prediction model. Complex machine learning algorithms would be expected to achieve more substantial gains in the case of multiple continuous predictors where non-linear relationships could be modelled more flexibly.



Conclusion

Overall summary

In this Phase Two project, we used the MADIP-FFY data, including the AEDC developmental outcomes and the 36 key child-level disadvantage indicators across four disadvantage lenses (sociodemographic, health conditions, geographic and risk factors) and six priority population indicators from our Phase One work. There was only one difference in the priority population indicators in this report, which included an additional priority population indicator of ‘parental ancestry’. We conducted additional data analysis over 3 parts: **Part 1**- Associations between child-level disadvantage and priority population indicators with developmental vulnerability on each of the five specific developmental domains; **Part 2**- Comparison of child-level and area-level measures of disadvantage; and **Part 3**- Child-level disadvantage and priority population indicators for prediction of DV1.

In **Part 1**, we showed the distribution of 36 key child-level disadvantage indicators and seven priority population indicators and children’s developmental vulnerability on each of the five specific domains, with the top two indicators presented and the full results in the appendices. We also estimated univariable associations of child disadvantage and priority population indicators with children’s **developmental vulnerability on the five specific domains** (PHY-WELL, EMO, SOC, LANG-COG, and COM-KNO). Overall, children who experienced disadvantage had higher risk of developmental vulnerability on all five specific domains than those not experiencing disadvantage. There was considerable overlap in the top 15 child disadvantage indicators between each of the five developmental domains, and a total of 18 child-level disadvantage and priority population indicators were identified across these five developmental domains. The majority of the indicators identified belonged to the sociodemographic lens, while the risk factor indicator ‘child not regularly read to at home’ had the strongest association with developmental vulnerability on all five domains. The indicators of ‘Child Care Subsidy income thresholds’ and ‘parental highest education’ had the next strongest associations with developmental vulnerability on the PHY-WELL, LANG-COG and COM-KNO domains, while ‘child has had any mental health issues’ and ‘Child Care Subsidy income thresholds’ had the next strongest univariable associations with developmental vulnerability on the EMO and SOC domains.

In **Part 2**, we compared the child-level measures of disadvantage (i.e., the 36 key child-level disadvantage indicators used in Phase One) with an **area-level measure** of disadvantage (i.e., IRSD-SEIFA, a commonly used measure of socioeconomic position). We found large differences in the distribution of the child-level disadvantage indicators between the SEIFA quintiles. Overall, there were higher proportions of children exposed to child-level disadvantage among those living in the most disadvantaged communities than those living in less disadvantaged communities. We also compared the univariable associations of each of the child-level and area-level disadvantage indicators with children’s DV1 and found that some child-level disadvantage indicators (e.g., lower household income, less home reading) had stronger associations than the area-level SEIFA quintile measure. Our findings indicate that children experiencing a particular aspect of child-level disadvantage were often missed if using the area-based SEIFA variable as a single proxy measure of child disadvantage. There is growing concern regarding the use of area-level SEIFA as a proxy measure of child disadvantage, given the large differences that exist within the same SEIFA quintile. Measuring child disadvantage and further understanding developmental vulnerability requires careful consideration of measures.

In **Part 3**, we built upon the univariable analyses completed in Phase One to develop a multivariable prediction model consisting of the **best combination** of child-level disadvantage indicators for identifying children most at risk of developmental vulnerability. After careful consideration of a series of candidate models, we determined that the parsimonious model comprised of 10 child-level disadvantage indicators – two indicators within each of the four disadvantage lenses and two indicators representing priority populations – was the optimal model for predicting DV1 based on the least number of predictors and high prediction accuracy. The two **sociodemographic** indicators identified were household income (using the childcare subsidy income thresholds) and parental highest education level. The two **health conditions** indicators were child’s chronic health issues and child’s mental health issues. The two **geographic**



indicators were house overcrowding and rented tenure type. The top two **risk factors** indicators were child not regularly read to at home and child born to a teenage mother. The two **priority population** indicators were child not proficient in English and parental ancestry. The optimal model had a specificity of 98% and sensitivity of 27%, corresponding to an overall predictive accuracy of 85%. Although the 10 indicators identified in the optimal set represent the aspects of disadvantage most likely to identify children at risk of DV1 in this data, a prediction model based on these disadvantage indicators alone is not sufficient to accurately quantify the risk of developmental vulnerability in children in their first year of school given the low sensitivity. Importantly, our analysis highlights that prediction of developmental vulnerability among children warrants the use of a diverse range of disadvantage indicators from different lenses, given the complex multi-faceted nature of child disadvantage and the consideration of other important factors outside of the measured aspects of disadvantage that are important in determining children's developmental vulnerability.

Data gaps and challenges

A range of data limitations and challenges should be noted.

Measuring child-level disadvantage over time

For many of the child-level disadvantage indicators, we were limited to measuring child-level disadvantage indicators at a single point in time due to a lack of **longitudinal data**. For example, data on parental highest education and whether a child was regularly read to at home were available only from the 2018 AEDC, which was collected the same time at which developmental vulnerability was assessed. Therefore, the associations between these indicators and developmental vulnerability were cross-sectional. It would be valuable to examine the longitudinal relationship between child-level disadvantage indicators and developmental vulnerability from a lifecourse perspective^{34,35} in order to capture the cumulative impact of exposure to disadvantage in early childhood prior to school entry.

Lack of specificity in child-level disadvantage

Due to limitations in the data available, we focused on **crude indicators** (e.g., binary, categorical) of child-level disadvantage, which indicated whether a child was exposed to a particular type of disadvantage or not. However, these data did not capture any information on chronicity, severity or frequency. Previous studies have shown that the more severe health conditions (e.g., child's mental health issues) children have, the more negative impacts they have on health and development.^{19,36}

Ethnicity and race

In this report, we measured 'parental ancestry' using data from the 2016 Census. However, it should be noted that we included 'Australian' as parts of 'Anglo-Celtic' groups due to limited data in the Census. However, Australian is a nationality and not an **ancestry or ethnicity**. It is somewhat problematic that 'Australian' is often used as a proxy for being White/Anglo-Celtic, which is exclusionary of those who also identify as Australian and are from other backgrounds.

We were unable to measure '**child ancestry**' using data from the 2016 Census, as it was challenging to correctly identify the child who was linked to 2018 AEDC. It was difficult to classify relationships between children and parents using Census data, as it does not include information on who the adults living in the same household are in relation to a child.



Recommendations

The following steps are recommended for enhancing the measurement of child disadvantage and developmental vulnerability in early childhood data collections.

Research

We focused on the whole population of children in 2018 AEDC for all analyses in this report. However, as shown in the literature, large disparities exist between some priority population subgroups (e.g., children from an Aboriginal or Torres Strait Islander background, children from culturally and linguistically diverse backgrounds or refugee backgrounds) and those not. It would be valuable in future to examine the distribution of child-level disadvantage indicators and their relationships with developmental vulnerability in these **subgroup populations**.

While children being from an Aboriginal or Torres Strait Islander background and other priority population indicators were not the strongest predictors of DV1 in this study, there is overwhelming evidence that the Indigenous population faces the greatest inequities across all indicators. Thus, the priority population indicator findings should be interpreted with caution. We examined the distribution of separate child-level disadvantage or priority population indicators with children's developmental vulnerability outcomes. It would be important to look at the **interplay** between child-level disadvantage indicators and priority population indicators together and evaluate how they may contribute to children's developmental vulnerability.^{37,38}

Future research may consider monitoring the distribution of child-level disadvantage indicators over time and examining the relationship between trajectories of child-level disadvantage and children's developmental vulnerability. The developmental **trajectories** approach can help to map the developmental course of child disadvantage, inform the role of cumulative disadvantage, and assess heterogeneity in response to interventions.

It would be worthwhile to examine and compare how a model with multiple indicators from one lens (e.g., sociodemographic) would perform relative to models with indicators across each of the lenses to predict children who are at the highest risk of developmental vulnerability. This would help to check the validity of the **CCC multidimensional framework** of child disadvantage² and inform the target population for interventions.

Causal modelling would be informative to specify and quantify causal pathways between disadvantage indicators and children's developmental vulnerability more precisely. These methods can be used to model the extent to which intervening on different combinations of modifiable factors (e.g., preschool attendance, family income, maternal education, housing, mental health) could potentially reduce inequities in children's developmental vulnerability. This in turn would help to further understand the **key drivers/mechanisms of developmental inequities** that can be leveraged for more precise policy interventions.

Data and practice

It is important to consider how to refine and enhance measurement of current indicators in existing data collections. For example, for the child disadvantage indicators within the geographic lens, we were limited to indicators at the household level. It may be of interest to explore **geographic indicators** beyond the household level, such as transport to workplace in the Census.³⁹

Housing plays a significant role in shaping a child's health and development.⁴⁰ In this report, we examined house crowding, dwelling type, tenure type, and residential mobility in the last five years. We were limited to the 2016 Census data, which only had data on the amount of rent paid by household for occupied private dwellings. Housing stress is defined as lower-income households (lowest 40% of income) that spend more than 30% of gross income on housing costs. Future data collections may consider **other housing variables** e.g., housing costs not only include rent payments, but also rate payments (e.g., water, council) and housing related mortgage payments.



Currently, data on ethnicity is not routinely collected in Australia, except for Aboriginal and Torres Strait Islander status. Due to the complexity of the Census data, it is challenging to derive the indicator of 'child ancestry'. It would be valuable to examine children's developmental vulnerability by **child ancestry** through future early childhood data collections. Further, categorising 'Australian' as parts of 'Anglo-Celtic' groups is somewhat problematic because Australian is a nationality and not an ancestry or ethnicity.^{13,41} Development of an ancestry indicator that more **accurately captures people's diverse backgrounds** needs to be prioritised.

For parental highest education, we used data from 2018 AEDC, which was collected at the same time as the outcome assessment. While data on parental highest education are also available from 2016 Census (i.e., at an earlier time before the outcome assessment), we were unable to identify which adult was the child's parents or primary carer. Future research may consider using more recently linked data with AEDC to indicate a more proximal source of **parental highest education**, such as data linkage with the Australian Census Longitudinal Dataset (ACL), to better elucidate the relationships between parental education and child disadvantage.

The indicator 'unpaid childcare' was derived from 2016 Census. In this report, we found that children with parents that were provided unpaid childcare had lower rates of developmental vulnerability across five AEDC domains, suggesting unpaid childcare is a protective factor of developmental vulnerability. However, the role of unpaid childcare is complex and should be viewed in conjunction with **other family factors** such as single-income families, extended families, and lack of provision of paid childcare.^{42,43}

Policy

While a range of child-level disadvantage indicators have been collected and used in government-based administrative datasets (e.g., AEDC, Census, MBS/PBS), they have some limitations and challenges which may be difficult to implement into practice. When making informed policy decisions regarding the use of these indicators identified in this work, governments should interpret the findings with caution and acknowledge the **measurement limitations**.

Governments need to consider the **language** that is used and **not stigmatise children** when reporting on a negative outcome. For example, children experiencing disadvantage rather than disadvantaged children.

Together our findings confirm that childhood disadvantage has a lasting negative impact on children's health and development. Addressing childhood disadvantage should consider a **social determinants framework** as there is no one 'silver bullet' policy, as part of the Early Years Strategy,⁴⁴ that eliminates developmental inequities in children.

Utilising government-based administrative data is increasingly gaining attention to build policy-relevant evidence. The MADIP-FFY data is a **powerful resource**. Governments should consider how to make this data more widely available to researchers and critical thinkers. This would help to encourage **collaboration** and ensure that **the maximum impact** is achieved to inform policy decision making.



References

- Braveman P, Barclay C. Health disparities beginning in childhood: A life-course perspective. *Pediatrics*. 2009;124(Suppl 3):S163-75. doi:<https://doi.org/10.1542/peds.2009-1100D>
- Goldfeld S, O'Connor M, Chong S, et al. The impact of multidimensional disadvantage over childhood on developmental outcomes in Australia. *International Journal of Epidemiology*. 2018;47(5):1485-1496. doi:<https://doi.org/10.1093/ije/dyy087>
- Heckman J, Mosso S. The economics of human development and social mobility. *Annual Review of Economics*. 2014;6:689-733. doi:<https://doi.org/10.1146/annurev-economics-080213-040753>
- Koh H, Oppenheimer S, Massin-Short S, Emmons K, Geller A, Viswanath K. Translating research evidence into practice to reduce health disparities: A social determinants approach. *American Journal of Public Health*. 2010;1(100):S72-80. <https://doi.org/10.2105/AJPH.2009.167353>
- Goldfeld S, O'Connor E, Sung V, et al. Potential indirect impacts of the COVID-19 pandemic on children: A narrative review using a community child health lens. *Medical Journal of Australia*. 2022;216(7):364-372. doi:<https://doi.org/10.5694/mja2.51368>
- Commission on Social Determinants of Health. *Closing the gap in a generation: Health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health*. 2008. <https://www.who.int/publications/i/item/WHO-IER-CSDH-08.1>
- Goldfeld S, Gray S, Azpitarte F, et al. Driving precision policy responses to child health and developmental inequities. *Health Equity*. 2019;3(1):489-494. doi:<https://doi.org/10.1089/heap.2019.0045>
- Goldfeld S, O'Connor M, Cloney D, et al. Understanding child disadvantage from a social determinants perspective. *Journal of Epidemiology & Community Health*. 2018;72(3):223-229. doi:<http://dx.doi.org/10.1136/jech-2017-209036>
- Heckman J, Masterov D. The productivity argument for investing in young children. *Review of Agricultural Economics*. 2007;29:446-493. doi:<https://doi.org/10.1111/j.1467-9353.2007.00359.x>
- Katz I, Redmond G. Investment in early childhood in Australia: International comparisons and recent trends. *Health Sociology Review*. 2009;18(1):94-107. doi:<https://doi.org/10.5172/hesr.18.1.94>
- Gray S, Pham C, Guo S, et al. *Measuring vulnerability and disadvantage in early childhood data collections*. 2023. <https://doi.org/10.25374/MCRI.21974780>
- Pink B. *Information paper: an introduction to socio-economic indexes for areas (SEIFA)*, 2006. 2008.
- Priest N, Alam O, Dunn K, et al. *Evaluation of the Speak Out Against Racism (SOAR) program pilot. CSRM working paper no. 2/2020*. 2020. 2020.
- Brinkman S. The predictive validity of the AEDI: Predicting later cognitive and behavioural outcomes. presented at: Australian Council for Educational Research Conference; 2014; Adelaide, Australia.
- Goldfeld S, O'Connor E, O'Connor M, et al. The role of preschool in promoting children's healthy development: Evidence from an Australian population cohort. *Early Childhood Research Quarterly*. 2016;35:40-48. doi:<http://dx.doi.org/10.1016/j.ecresq.2015.11.001>
- O'Connor M, O'Connor E, Gray S, Goldfeld S. Trends in preschool attendance in Australia following major policy reform: Updated evidence six years following a commitment to universal access. *Early Childhood Research Quarterly*. 2020;51:93-99. doi:<https://doi.org/10.1016/j.ecresq.2019.09.002>
- CCCH, TICHR. *A snapshot of early childhood development in Australia: AEDI National Report 2009*. 2009.
- Martinez W, Carter JS, Legato LJ. Social Competence in Children with Chronic Illness: A Meta-analytic Review. Article. *Journal of Pediatric Psychology*. 2011;36(8):878-890. doi:10.1093/jpepsy/jsr035
- Fardell JE, Hu N, Wakefield CE, et al. Impact of hospitalizations due to chronic health conditions on early child development. *Journal of pediatric psychology*. 2023;jsad025.
- Australian Government Department of Health and Ageing. KidsMatter. Australian Government Department of Health and Ageing. Accessed Mar 10, 2017. <http://www.kidsmatter.edu.au/>
- Singh G. Mental health of adolescents in relation to emotional maturity and parent child relationship. *Voice of Research*. 2014;3(2):4-6.
- Australian Bureau of Statistics (2021). *Census of Population and Housing: Census dictionary. Tenure type (TEND)*. <https://www.abs.gov.au/census/guide-census-data/census-dictionary/2021/variables-topic/housing/tenure-type-tend>
- Goldfeld S, O'Connor M, Mithen J, Sayers M, Brinkman S. Early development of emerging and English-proficient bilingual children at school entry in an Australian population cohort. *International Journal of Behavioral Development*. 2014;38:42-51. doi:<http://dx.doi.org/10.1177/0165025413505945>
- Whiteside KE, Gooch D, Norbury CF. English language proficiency and early school attainment among children learning English as an additional language. *Child development*. 2017;88(3):812-827.
- Lim P, Gemici S. *Measuring the socioeconomic status of Australian youth*. National Centre for Vocational Education Research; 2011.
- Australian Bureau of Statistics. *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2011*. 2011. 2033.0.55.001. <http://www.abs.gov.au/ausstats/abs@/nsf/Lookup/2033.0.55.001main+features100052011>
- Australian Bureau of Statistics (2021). *Socio-Economic Indexes for Areas (SEIFA), Australia*. <https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release>
- Bendre MR, Thool VR. Analytics, challenges and applications in big data environment: a survey. *Journal of Management Analytics*. 2016;3(3):206-239.
- Hao J. Supervised Machine Learning. *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment: With Examples in R and Python*. 2021:159-171.
- Dormann CF, Elith J, Bacher S, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013;36:27-46.
- Lieberman M, Morris J. The Precise Effect of Multicollinearity on Classification Prediction. 01/01 2014;40:5-10.



32. Brownlee J. A gentle introduction to threshold-moving for imbalanced classification. *Machine Learning Mastery*. 2020;
33. Laan MJvd, Polley EC, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology*. 2007;6(1) doi:<https://doi.org/10.2202/1544-6115.1309>
34. Graham H. Building an inter-disciplinary science of health inequalities: the example of lifecourse research. *Social Science & Medicine*. 12// 2002;55(11):2005-2016. doi:[http://dx.doi.org/10.1016/S0277-9536\(01\)00343-4](http://dx.doi.org/10.1016/S0277-9536(01)00343-4)
35. Graham H, Power C. *Childhood disadvantage and adult health: A lifecourse framework*. Health Development Agency; 2004.
36. Kivimäki M, Batty GD, Pentti J, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. *The Lancet Public Health*. 2020;5(3):e140-e149. doi:10.1016/S2468-2667(19)30248-8
37. Bronfenbrenner U. Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*. 1986;22:723-742.
38. Earls F, Carlson M. The social ecology of child health and wellbeing. *Annual Review of Public Health*. 2001;22:143-166.
39. Villanueva K, Alderton A, Higgs C, Badland H, Goldfeld S. Data to Decisions: Methods to Create Neighbourhood Built Environment Indicators Relevant for Early Childhood Development. *Int J Environ Res Public Health*. May 3 2022;19(9)doi:10.3390/ijerph19095549
40. Dunn JR. Housing and Healthy Child Development: Known and Potential Impacts of Interventions. *Annual Review of Public Health*. 2020/04/02 2020;41(1):381-396. doi:10.1146/annurev-publhealth-040119-094050
41. Priest N, Chong S, Truong M, et al. *Findings from the 2017 Speak Out Against Racism (SOAR) student and staff surveys*. CSRM working paper no. 3/2019. 2019. https://csrm.cass.anu.edu.au/sites/default/files/docs/2019/8/CSRM-WP-SOAR_PUBLISH_1.pdf
42. Nakamichi K, Takahashi M, Sunagami F, Iwata M. The relationship between child-centered teaching attitudes in childcare centers and the socio-emotional development of Japanese toddlers. *Early Childhood Research Quarterly*. 2022/04/01/ 2022;59:162-171. doi:<https://doi.org/10.1016/j.ecresq.2021.11.014>
43. Cook K, Corr L, Breitzkreuz R. The framing of Australian childcare policy problems and their solutions. Article. *Critical Social Policy*. 2017;37(1):42-63. doi:10.1177/0261018316653952
44. Australian Labor Party. Early Years Strategy. Australian Labor Party. 30 August, 2022. <https://www.alp.org.au/policies/early-years-strategy>



Appendices

Appendix A. Part 1: Developmental vulnerability on five specific domains

Table A.1. The percentage of children's developmental vulnerability on five specific domains by the key child disadvantage indicators.

Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Sociodemographic					
Poverty line:					
Above poverty line	9.1% (20,721)	8.0% (18,126)	9.1% (20,707)	5.8% (13,261)	7.0% (16,014)
Poverty line or below	11.3% (4,095)	9.6% (3,460)	11.6% (4,202)	8.6% (3,114)	10.6% (3,833)
Family eligible for a Low Income Health Care Card:					
No	7.0% (11,847)	6.7% (11,268)	7.3% (12,250)	3.7% (6,161)	5.0% (8,517)
Yes	13.7% (12,922)	10.9% (10,271)	13.4% (12,613)	10.8% (10,185)	12.0% (11,292)
Family Tax Benefit A, based on income group:					
Greater than \$99,864	6.3% (8,460)	6.2% (8,320)	6.7% (8,899)	3.0% (3,956)	4.2% (5,599)
\$56,137 to \$99,864	10.8% (6,323)	9.0% (5,263)	10.6% (6,230)	7.4% (4,333)	9.3% (5,483)
\$56,137 or less	14.1% (10,056)	11.3% (8,019)	13.7% (9,797)	11.4% (8,096)	12.3% (8,781)
Family Tax Benefit B, based on income group:					
Greater than \$100,900	6.3% (8,341)	6.2% (8,216)	6.6% (8,779)	2.9% (3,890)	4.2% (5,501)
\$100,900 or less	12.6% (16,498)	10.2% (13,386)	12.3% (16,147)	9.5% (12,495)	10.9% (14,362)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Child Care Subsidy income thresholds (4 categories):					
Greater than \$254,305	4.6% (809)	5.0% (876)	5.1% (890)	1.4% (250)	2.2% (383)
Greater than \$175,015 to \$254,305	5.8% (1,914)	5.9% (1,961)	6.0% (2,000)	2.4% (781)	3.2% (1,052)
Greater than \$70,015 to \$175,015	7.9% (9,678)	7.2% (8,866)	8.0% (9,898)	4.5% (5,521)	6.2% (7,581)
\$0 to \$70,015	13.9% (12,434)	11.1% (9,899)	13.5% (12,140)	11.0% (9,833)	12.1% (10,851)
Child Care Subsidy income thresholds (2 categories):					
Greater than \$70,015	7.1% (12,400)	6.8% (11,703)	7.4% (12,786)	3.8% (6,552)	5.2% (9,012)
\$70,015 or less	13.9% (12,434)	11.1% (9,899)	13.5% (12,140)	11.0% (9,833)	12.1% (10,851)
Parental highest education level:					
Bachelor's degree or above	6.0% (7,158)	5.9% (6,987)	6.3% (7,454)	2.4% (2,854)	4.5% (5,286)
Advanced Diploma or Diploma	9.3% (4,459)	8.1% (3,913)	9.6% (4,637)	6.0% (2,875)	7.3% (3,496)
Certificate level I to IV ^a	11.4% (6,476)	10.2% (5,793)	11.5% (6,547)	8.0% (4,560)	8.5% (4,865)
Year 12 or below	16.9% (6,046)	12.2% (4,375)	15.3% (5,502)	15.0% (5,357)	15.5% (5,553)
Parent highest occupation:					
Managers / Professionals	5.9% (6,359)	5.9% (6,397)	6.1% (6,616)	2.4% (2,600)	3.6% (3,928)
Technicians / Other types of workers ^b	8.2% (4,957)	8.0% (4,764)	9.0% (5,421)	5.5% (3,278)	7.1% (4,280)
Labourers / Others ^c	10.1% (3,712)	8.6% (3,168)	10.4% (3,822)	6.7% (2,468)	9.1% (3,361)
Parent highest occupation:					
White collar	6.7% (10,134)	6.7% (10,055)	7.1% (10,717)	3.2% (4,893)	4.5% (6,825)
Blue collar	9.0% (4,891)	7.9% (4,274)	9.5% (5,140)	6.4% (3,453)	8.8% (4,744)
Parent employment status:					
Employed	7.3% (14,536)	7.0% (13,924)	7.7% (15,336)	4.0% (8,071)	5.6% (11,134)
Not employed	12.5% (1,956)	10.4% (1,623)	12.8% (2,010)	9.0% (1,401)	11.3% (1,762)
Parent employment average duration:					
Greater than 4 years	7.4% (13,975)	7.1% (13,419)	7.8% (14,705)	4.1% (7,655)	5.4% (10,244)
4 years or less	14.4% (11,470)	11.1% (8,771)	13.6% (10,866)	11.5% (9,150)	12.6% (10,055)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Parent received social support payment:					
Age pension support payment:					
No	9.4% (25,739)	8.2% (22,432)	9.5% (25,869)	6.2% (17,005)	7.5% (20,561)
Yes	15.9% (77)	20.1% (97)	20.9% (101)	12.4% (60)	13.5% (65)
Carer support payment:					
No	8.9% (22,525)	7.9% (20,033)	9.0% (22,930)	5.7% (14,580)	7.1% (18,044)
Yes	17.0% (3,291)	13.0% (2,493)	15.8% (3,046)	12.9% (2,485)	13.4% (2,583)
Rent assistance support payment:					
No	7.0% (12,067)	6.5% (11,213)	7.2% (12,310)	4.0% (6,902)	5.2% (9,012)
Yes	13.6% (13,749)	11.2% (11,313)	13.5% (13,660)	10.0% (10,163)	11.5% (11,613)
Family support payment:					
No	5.7% (1,021)	5.3% (950)	5.9% (1,061)	3.0% (530)	4.5% (803)
Yes	9.7% (24,800)	8.5% (21,576)	9.7% (24,912)	6.5% (16,535)	7.8% (19,825)
Employment support payment:					
No	8.4% (19,807)	7.4% (17,402)	8.4% (19,917)	5.2% (12,335)	6.5% (15,493)
Yes	16.4% (6,018)	14.0% (5,124)	16.5% (6,061)	12.9% (4,730)	14.0% (5,132)
Student support payment:					
No	9.4% (25,033)	8.2% (21,798)	9.4% (25,177)	6.2% (16,504)	7.5% (20,003)
Yes	13.4% (784)	12.5% (728)	13.6% (794)	9.8% (570)	10.7% (624)
Disability support payment:					
No	9.1% (24,371)	8.1% (21,446)	9.3% (24,694)	6.0% (15,978)	7.3% (19,556)
Yes	21.7% (1,446)	16.3% (1,080)	19.2% (1,277)	16.4% (1,087)	16.1% (1,069)
Any type of social security payments:					
No	5.6% (959)	5.2% (903)	5.8% (997)	2.80% (479)	4.4% (752)
Yes	9.7% (24,862)	8.5% (21,623)	9.7% (24,976)	6.5% (16,586)	7.8% (19,876)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Parent received special childcare benefit:					
At risk childcare benefit:					
No	8.7% (17,430)	8.4% (16,644)	9.4% (18,833)	5.4% (10,722)	6.8% (13,548)
Yes	23.5% (1,352)	21.1% (1,207)	23.8% (1,368)	16.7% (959)	14.6% (840)
Financial hardship childcare benefit:					
No	8.8% (17,367)	8.4% (16,524)	9.5% (18,705)	5.5% (10,761)	6.8% (13,472)
Yes	16.4% (1,415)	15.4% (1,327)	17.3% (1,496)	10.7% (923)	10.6% (916)
Grandparent childcare benefit:					
No	9.1% (18,505)	8.6% (17,574)	9.7% (19,877)	5.6% (11,459)	7.0% (14,199)
Yes	24.6% (277)	24.9% (277)	28.8% (324)	19.2% (216)	16.7% (188)
Jobs education and training childcare benefit:					
No	8.9% (17,495)	8.4% (16,518)	9.5% (18,738)	5.5% (10,778)	6.9% (13,510)
Yes	14.8% (1,288)	15.3% (1,333)	16.8% (1,464)	10.4% (906)	10.1% (881)
Any special childcare benefit payments:					
No	8.3% (15,272)	7.9% (14,551)	8.9% (16,485)	5.0% (9,241)	6.5% (12,070)
Yes	17.1% (3,511)	16.1% (3,300)	18.1% (3,717)	11.9% (2,434)	11.3% (2,321)
Child with a lone parent family:					
No	7.8% (15,526)	7.0% (13,847)	7.9% (15,821)	4.7% (9,357)	6.4% (12,857)
Yes	14.2% (9,567)	11.9% (7,970)	13.9% (9,365)	10.7% (7,209)	10.8% (7,236)
Household size with 6 or more people:					
5 people or less	9.4% (23,486)	8.3% (20,763)	9.5% (23,828)	6.2% (15,430)	7.5% (18,725)
6 people or more	10.2% (1,387)	6.4% (875)	8.3% (1,138)	7.2% (979)	8.6% (1,170)
Health conditions					
Parent has had any chronic health issue(s):					
No	8.6% (15,713)	7.7% (13,991)	8.6% (15,744)	5.7% (10,316)	6.8% (12,479)
Yes	11.1% (10,021)	9.4% (8,472)	11.2% (10,155)	7.4% (6,660)	8.9% (8,043)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Child has had any chronic health issue(s):					
No	8.7% (20,819)	7.5% (18,005)	8.7% (20,786)	5.8% (13,917)	6.7% (16,108)
Yes	14.8% (4,989)	13.5% (4,511)	15.4% (5,173)	9.3% (3,134)	13.4% (4,502)
Parent has had any mental health issue(s):					
No	7.8% (8,803)	6.5% (7,254)	7.9% (8,891)	5.5% (6,180)	7.7% (8,702)
Yes	10.5% (16,931)	9.5% (15,209)	10.6% (16,999)	6.7% (10,796)	7.4% (11,820)
Parent mental health issue duration:					
Parent has no mental health issues or has had a mental health issue less than one year	8.1% (13,253)	6.9% (11,283)	8.3% (13,453)	5.6% (9,180)	7.6% (12,446)
Parent has had a mental health issue greater than one year	11.3% (12,481)	10.2% (11,180)	11.3% (12,437)	7.1% (7,796)	7.3% (8,076)
Child has had any mental health issue(s):					
No	8.9% (22,477)	7.1% (17,975)	8.5% (21,512)	6.0% (15,111)	7.3% (18,445)
Yes	15.8% (3,330)	21.7% (4,541)	21.2% (4,445)	9.2% (1,940)	10.3% (2,165)
Child mental health issue duration:					
Child has no mental health issues or has had a mental health issue less than one year	9.3% (25,010)	8.0% (21,316)	9.2% (24,812)	6.2% (16,616)	7.5% (20,131)
Child has had a mental health issue greater than one year	16.0% (795)	24.3% (1,200)	23.0% (1,144)	8.8% (435)	9.8% (486)
Geographic					
House crowding (3 or more additional bedrooms needed):					
No	8.4% (18,847)	7.6% (16,947)	8.6% (19,220)	5.1% (11,457)	6.6% (14,801)
Yes	19.4% (241)	12.2% (152)	17.2% (214)	19.4% (241)	21.8% (271)
House crowding (1 or more additional bedrooms needed):					
No	8.0% (16,405)	7.5% (15,235)	8.3% (16,929)	4.6% (9,481)	6.0% (12,314)
Yes	13.4% (2,687)	9.3% (1,857)	12.5% (2,511)	11.0% (2,213)	13.8% (2,767)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Dwelling type:					
Private dwellings	8.6% (19,915)	7.7% (17,822)	8.8% (20,283)	5.3% (12,354)	6.8% (15,808)
Collective dwellings	11.1% (72)	6.3% (41)	9.2% (60)	5.7% (37)	6.9% (45)
Tenure type:					
Own	6.6% (9,753)	6.4% (9,354)	7.0% (10,231)	3.5% (5,105)	4.9% (7,224)
Rented ^d	12.0% (9,751)	10.1% (8,147)	11.9% (9,661)	8.5% (6,913)	10.1% (8,179)
Child has moved residence address in the last 5 years					
No	7.8% (6,515)	7.0% (5,795)	7.8% (6,496)	4.7% (3,965)	5.7% (4,750)
Yes	8.9% (12,644)	8.0% (11,386)	9.1% (12,991)	5.5% (7,775)	7.3% (10,385)
Risk factors					
Preschool non-attendance:					
No	8.6% (20,981)	8.0% (19,502)	8.9% (21,629)	5.3% (12,966)	6.5% (15,921)
Yes	18.3% (3,279)	11.5% (2,056)	15.6% (2,794)	15.1% (2,706)	16.9% (3,036)
Childcare non-attendance:					
No	9.1% (18,782)	8.7% (17,851)	9.8% (20,201)	5.7% (11,675)	7.0% (14,387)
Yes	10.3% (7,034)	6.9% (4,675)	8.5% (5,769)	7.9% (5,390)	9.2% (6,238)
Unpaid childcare:					
No	13.8% (2,122)	11.1% (1,700)	14.3% (2,196)	12.5% (1,911)	14.1% (2,165)
Yes	8.5% (19,560)	7.7% (17,703)	8.6% (19,883)	5.1% (11,759)	6.5% (15,089)
Child's age group at childcare entry:					
0-2 years	8.9% (15,389)	8.9% (15,277)	9.8% (16,932)	5.4% (9,310)	6.5% (11,159)
3-6 years	10.3% (3,397)	7.8% (2,574)	9.9% (3,276)	7.2% (2,365)	9.8% (3,233)
Child is not regularly read to at home:					
No	7.0% (17,330)	6.6% (16,389)	7.2% (17,972)	3.7% (9,230)	5.4% (13,357)
Yes	41.2% (6,976)	28.3% (4,768)	36.9% (6,248)	37.7% (6,366)	33.4% (5,655)



Indicator	PHY-WELL % (n)	EMO % (n)	SOC % (n)	LANG-COG % (n)	COM-KNO % (n)
Mother's age at birth (teenager):					
20 years or older	9.1% (23,839)	8.0% (20,833)	9.1% (23,905)	5.9% (15,378)	7.2% (18,863)
Younger than 20 years	19.1% (1,591)	16.7% (1,381)	19.6% (1,636)	16.7% (1,390)	15.8% (1,315)
Mother's age at birth (later):					
Younger than 35 years	9.9% (19,013)	8.7% (16,625)	10.0% (19,293)	6.8% (13,086)	8.1% (15,495)
35 years or older	8.3% (6,417)	7.2% (5,589)	8.1% (6,244)	4.8% (3,682)	6.0% (4,680)
Parental death:					
No	9.4% (25,718)	8.3% (22,436)	9.5% (25,858)	6.2% (16,999)	7.5% (20,560)
Yes	16.5% (98)	15.1% (90)	18.8% (112)	11.1% (66)	10.9% (65)

^a Certificate level I to IV including trade qualification. ^b Technicians, trade workers, community and personal service workers, and sales workers. ^c Clerical and administrative workers, machinery operators and driver, and labourers.

^d The tenure category 'living in residences that were rented' also includes 'occupied rent free', 'occupied under a life tenure scheme', and 'other tenure type'. *Abbreviations:* AEDC, Australian Early Development Census; COM-KNO Communication skills and general knowledge; EMO Emotional maturity; LANG-COG Language and cognitive skills (school-based); PHY-WELL Physical health and wellbeing; SOC Social competence.



Table A.2. Associations between child disadvantage indicators across the four social determinant lenses and developmental vulnerability on five specific domains.

Indicator	PHY-WELL RR (95% CI)	EMO RR (95% CI)	SOC RR (95% CI)	LANG-COG RR (95% CI)	COM-KNO RR (95% CI)
Sociodemographic					
Poverty line:					
Above poverty line	Reference	Reference	Reference	Reference	Reference
Poverty line or below	1.24 (1.21, 1.28)	1.20 (1.16, 1.24)	1.28 (1.24, 1.32)	1.48 (1.42, 1.53)	1.51 (1.46, 1.56)
Family eligible for a Low Income Health Care Card	1.95 (1.91, 2.00)	1.63 (1.59, 1.68)	1.84 (1.80, 1.89)	2.96 (2.87, 3.06)	2.38 (2.31, 2.44)
Family Tax Benefit A, based on income group:					
Greater than \$99,864	Reference	Reference	Reference	Reference	Reference
\$56,137 to \$99,864	1.70 (1.65, 1.76)	1.44 (1.39, 1.49)	1.59 (1.55, 1.64)	2.49 (2.39, 2.60)	2.23 (2.15, 2.31)
\$56,137 or less	2.23 (2.17, 2.29)	1.81 (1.75, 1.86)	2.06 (2.01, 2.12)	3.84 (3.70, 3.98)	2.94 (2.84, 3.03)
Family Tax Benefit B, based on income group:					
Greater than \$100,900	Reference	Reference	Reference	Reference	Reference
\$100,900 or less	1.99 (1.94, 2.04)	1.64 (1.60, 1.68)	1.85 (1.80, 1.90)	3.23 (3.12, 3.35)	2.63 (2.55, 2.71)
Child Care Subsidy income thresholds (4 categories):					
Greater than \$254,305	Reference	Reference	Reference	Reference	Reference
Greater than \$175,015 to \$254,305	1.25 (1.16, 1.36)	1.19 (1.10, 1.28)	1.19 (1.10, 1.29)	1.66 (1.44, 1.91)	1.45 (1.29, 1.63)
Greater than \$70,015 to \$175,015	1.71 (1.59, 1.83)	1.45 (1.35, 1.55)	1.59 (1.49, 1.70)	3.16 (2.78, 3.58)	2.83 (2.56, 3.13)
\$0 to \$70,015	3.01 (2.81, 3.23)	2.22 (2.07, 2.37)	2.68 (2.51, 2.86)	7.72 (6.81, 8.74)	5.55 (5.02, 6.14)
Child Care Subsidy income thresholds (2 categories):					
Greater than \$70,015	Reference	Reference	Reference	Reference	Reference
\$70,015 or less	1.94 (1.90, 1.99)	1.64 (1.60, 1.68)	1.84 (1.80, 1.88)	2.91 (2.82, 3.00)	2.33 (2.27, 2.40)
Parental highest education level:					
Bachelor's degree or above	Reference	Reference	Reference	Reference	Reference
Advanced Diploma or Diploma	1.53 (1.48, 1.59)	1.38 (1.33, 1.43)	1.53 (1.48, 1.59)	2.48 (2.36, 2.61)	1.63 (1.56, 1.70)
Certificate level I to IV ^a	1.88 (1.82, 1.95)	1.73 (1.67, 1.79)	1.83 (1.77, 1.89)	3.33 (3.18, 3.49)	1.92 (1.85, 1.99)
Year 12 or below	2.80 (2.71, 2.89)	2.08 (2.00, 2.15)	2.44 (2.37, 2.53)	6.22 (5.95, 6.50)	3.48 (3.36, 3.61)



Indicator	PHY-WELL RR (95% CI)	EMO RR (95% CI)	SOC RR (95% CI)	LANG-COG RR (95% CI)	COM-KNO RR (95% CI)
Parent highest occupation:					
Managers / Professionals	Reference	Reference	Reference	Reference	Reference
Technicians / Other types of workers ^b	1.40 (1.35, 1.45)	1.34 (1.29, 1.39)	1.47 (1.42, 1.52)	2.26 (2.15, 2.38)	1.96 (1.88, 2.04)
Labourers / Others ^c	1.71 (1.64, 1.78)	1.45 (1.39, 1.51)	1.69 (1.63, 1.76)	2.78 (2.64, 2.93)	2.51 (2.40, 2.62)
Parent highest occupation:					
White collar	Reference	Reference	Reference	Reference	Reference
Blue collar	1.34 (1.3, 1.39)	1.18 (1.14, 1.23)	1.33 (1.29, 1.38)	1.96 (1.88, 2.05)	1.93 (1.87, 2.00)
Parent was not employed	1.72 (1.64, 1.8)	1.49 (1.42, 1.57)	1.68 (1.60, 1.75)	2.22 (2.10, 2.35)	2.02 (1.93, 2.12)
Parent employment average duration:					
Greater than 4 years	Reference	Reference	Reference	Reference	Reference
4 years or less	1.95 (1.90, 1.99)	1.55 (1.51, 1.59)	1.75 (1.71, 1.80)	2.84 (2.76, 2.92)	2.33 (2.27, 2.39)
Parent received social support payment:					
Age pension support payment	1.69 (1.38, 2.08)	2.37 (1.98, 2.85)	2.21 (1.85, 2.63)	1.99 (1.57, 2.53)	1.76 (1.40, 2.22)
Carer support payment	1.92 (1.86, 1.99)	1.64 (1.58, 1.70)	1.75 (1.69, 1.81)	2.24 (2.16, 2.34)	1.88 (1.81, 1.96)
Rent assistance support payment	1.93 (1.89, 1.98)	1.71 (1.67, 1.76)	1.88 (1.84, 1.93)	2.50 (2.43, 2.57)	2.19 (2.13, 2.24)
Family support payment	1.70 (1.60, 1.81)	1.59 (1.49, 1.69)	1.64 (1.55, 1.75)	2.18 (2.00, 2.37)	1.73 (1.62, 1.85)
Employment support payment	1.96 (1.91, 2.01)	1.90 (1.85, 1.96)	1.96 (1.91, 2.01)	2.48 (2.40, 2.55)	2.14 (2.07, 2.20)
Student support payment	1.43 (1.34, 1.53)	1.53 (1.43, 1.64)	1.44 (1.35, 1.54)	1.56 (1.44, 1.69)	1.42 (1.32, 1.54)
Disability support payment	2.38 (2.27, 2.49)	2.02 (1.91, 2.14)	2.07 (1.97, 2.18)	2.73 (2.58, 2.89)	2.19 (2.07, 2.32)
Any type of social security payments	1.76 (1.65, 1.87)	1.61 (1.51, 1.72)	1.69 (1.59, 1.80)	2.33 (2.13, 2.55)	1.79 (1.66, 1.92)
Parent received special childcare benefit:					
At risk childcare benefit	2.70 (2.57, 2.83)	2.52 (2.39, 2.66)	2.52 (2.41, 2.65)	3.09 (2.91, 3.28)	2.15 (2.02, 2.30)
Financial hardship childcare benefit	1.86 (1.77, 1.95)	1.83 (1.74, 1.93)	1.82 (1.74, 1.91)	1.94 (1.82, 2.07)	1.55 (1.45, 1.65)
Grandparent childcare benefit	2.72 (2.45, 3.01)	2.88 (2.60, 3.19)	2.96 (2.70, 3.25)	3.42 (3.03, 3.86)	2.40 (2.11, 2.74)
Jobs education and training childcare benefit	1.66 (1.57, 1.75)	1.82 (1.73, 1.92)	1.76 (1.68, 1.85)	1.88 (1.76, 2.00)	1.46 (1.37, 1.56)
Any special childcare benefit payments	2.07 (2.00, 2.14)	2.04 (1.97, 2.12)	2.03 (1.97, 2.10)	2.37 (2.28, 2.48)	1.73 (1.66, 1.80)



Indicator	PHY-WELL RR (95% CI)	EMO RR (95% CI)	SOC RR (95% CI)	LANG-COG RR (95% CI)	COM-KNO RR (95% CI)
Child with a lone parent family	1.83 (1.79, 1.87)	0.77 (0.72, 0.82)	0.87 (0.83, 0.93)	2.29 (2.22, 2.36)	1.67 (1.62, 1.72)
Household size with 6 or more people	1.08 (1.03, 1.14)	1.30 (1.22, 1.39)	1.15 (1.08, 1.20)	1.16 (1.09, 1.24)	1.14 (1.08, 1.21)
Health conditions					
Parent has had any chronic health issue(s)	1.29 (1.26, 1.32)	1.22 (1.19, 1.25)	1.30 (1.27, 1.33)	1.30 (1.26, 1.34)	1.30 (1.27, 1.34)
Child has had any chronic health issue(s)	1.70 (1.66, 1.75)	1.79 (1.73, 1.84)	1.77 (1.72, 1.82)	1.60 (1.55, 1.66)	1.99 (1.93, 2.05)
Parent has had any mental health issue(s)(MHI)	1.35 (1.31, 1.38)	1.47 (1.43, 1.51)	1.34 (1.31, 1.37)	1.22 (1.19, 1.26)	0.95 (0.93, 0.98)
Parent mental health issue duration:					
No MHI issues or had MHI for less than one year	Reference	Reference	Reference	Reference	Reference
Greater than one year	1.40 (1.36, 1.43)	1.47 (1.43, 1.50)	1.37 (1.34, 1.40)	1.26 (1.22, 1.30)	0.96 (0.94, 0.99)
Child has had any mental health issue(s)	1.78 (1.72, 1.84)	3.04 (2.95, 3.13)	2.48 (2.41, 2.55)	1.54 (1.47, 1.61)	1.41 (1.35, 1.47)
Child mental health issue duration:					
No MHI issues or had MHI for less than one year	Reference	Reference	Reference	Reference	Reference
Greater than one year	1.72 (1.61, 1.83)	3.04 (2.89, 3.20)	2.49 (2.36, 2.62)	1.42 (1.29, 1.55)	1.29 (1.18, 1.40)
Geographic					
House crowding (3 or more additional bedrooms needed)	2.30 (2.05, 2.58)	1.54 (1.32, 1.80)	1.99 (1.76, 2.26)	3.73 (3.32, 4.19)	3.29 (2.96, 3.66)
House crowding (1 or more additional bedrooms needed)	1.66 (1.60, 1.73)	1.24 (1.19, 1.30)	1.51 (1.45, 1.57)	2.38 (2.28, 2.49)	2.28 (2.19, 2.37)
Dwelling type:					
Private dwellings	Reference	Reference	Reference	Reference	Reference
Collective dwellings	1.29 (1.03, 1.60)	0.78 (0.57, 1.06)	1.05 (0.83, 1.34)	1.07 (0.78, 1.46)	0.99 (0.75, 1.32)
Tenure type:					
Own	Reference	Reference	Reference	Reference	Reference
Rented ^d	1.81 (1.76, 1.86)	1.58 (1.54, 1.63)	1.71 (1.67, 1.76)	2.45 (2.37, 2.54)	2.05 (1.99, 2.12)
Child has moved residence address in the last 5 years	1.14 (1.11, 1.17)	1.15 (1.12, 1.19)	1.18 (1.14, 1.21)	1.15 (1.11, 1.20)	1.28 (1.24, 1.33)



Indicator	PHY-WELL RR (95% CI)	EMO RR (95% CI)	SOC RR (95% CI)	LANG-COG RR (95% CI)	COM-KNO RR (95% CI)
Risk factors					
Preschool non-attendance	2.13 (2.06, 2.20)	1.44 (1.38, 1.50)	1.76 (1.70, 1.82)	2.84 (2.74, 2.95)	2.59 (2.50, 2.68)
Childcare non-attendance	1.13 (1.10, 1.16)	1.27 (1.22, 1.30)	1.16 (1.12, 1.19)	1.40 (1.35, 1.44)	1.31 (1.27, 1.35)
Child attended unpaid childcare	1.64 (1.56, 1.69)	1.45 (1.39, 1.52)	1.67 (1.59, 1.72)	2.44 (2.33, 2.56)	2.17 (2.08, 2.27)
Child's age group at childcare entry:					
0-2 years	Reference	Reference	Reference	Reference	Reference
3-6 years	1.15 (1.11, 1.19)	0.88 (0.85, 0.92)	1.01 (0.97, 1.05)	1.33 (1.27, 1.39)	1.51 (1.46, 1.57)
Child not regularly read to at home	5.92 (5.79, 6.06)	4.29 (4.17, 4.42)	5.12 (5.00, 5.24)	10.16 (9.88, 10.44)	6.22 (6.06, 6.39)
Mother's age at birth is 20 years or younger	2.09 (2.00, 2.19)	2.08 (1.98, 2.19)	2.14 (2.05, 2.24)	2.84 (2.70, 2.98)	2.18 (2.07, 2.30)
Mother's age at birth is younger than 35 years	1.19 (1.16, 1.22)	1.19 (1.16, 1.23)	1.25 (1.20, 1.28)	1.43 (1.39, 1.49)	1.33 (1.30, 1.37)
Parental death	1.75 (1.46, 2.10)	1.83 (1.51, 2.22)	1.99 (1.68, 2.35)	1.78 (1.42, 2.23)	1.45 (1.15, 1.82)

^a Certificate level I to IV including trade qualification. ^b Technicians, trade workers, community and personal service workers, and sales workers. ^c Clerical and administrative workers, machinery operators and driver, and labourers. ^d The tenure category 'living in residences that were rented' also includes 'occupied rent free', 'occupied under a life tenure scheme', and 'other tenure type'. *Abbreviations:* AEDC Australian Early Development Census; CI Confidence Interval; COM-KNO Communication skills and general knowledge; EMO Emotional maturity; LANG-COG Language and cognitive skills (school-based); PHY-WELL Physical health and wellbeing; RR Risk ratio; SOC Social competence.

**Table A.3.** The percentage of children's developmental vulnerability on five specific domains by the priority population indicators

Characteristics	PHY-WELL %(n)	EMO %(n)	SOC %(n)	LANG-COG %(n)	COM-KNO %(n)
Child country of birth:					
Australia	9.5% (24,677)	8.3% (21,537)	9.5% (24,677)	6.2% (16,154)	7.3% (19,009)
Other English-Speaking country	9.3% (354)	7.5% (286)	9.1% (347)	6.7% (256)	7.7% (295)
Other Non-English speaking country	8.7% (750)	8.0% (682)	10.5% (906)	7.1% (607)	14.7% (1,263)
Parent country of birth:					
Australia	10.4% (17,176)	9.1% (15,029)	10.0% (16,564)	6.8% (11,237)	6.5% (10,853)
Other English-Speaking country	8.0% (2,979)	7.4% (2,735)	8.0% (2,989)	4.7% (1,730)	6.0% (2,213)
Other Non-English speaking country	8.0% (5,655)	6.8% (4,757)	9.1% (6,406)	5.8% (4,083)	10.7% (7,544)
Child Aboriginal and Torres Strait Islander status					
No	8.7% (22,425)	7.8% (19,945)	8.9% (22,794)	5.4% (13,881)	6.9% (17,711)
Yes	20.4% (3,385)	15.6% (2,578)	19.2% (3,179)	19.2% (3,179)	17.6% (2,914)
Parent ancestry:					
Anglo-Celtic ^a	10.4% (3,000)	8.9% (2,572)	9.8% (2,831)	6.6% (1,910)	6.7% (1,939)
Aboriginal and/or Torres Strait Islander	22.1% (538)	18.1% (439)	21.2% (516)	23.5% (570)	21.2% (516)
Pacific Islander/Māori	12.3% (860)	8.6% (600)	11.1% (777)	11.1% (774)	12.4% (864)
Middle Eastern	10.6% (1,264)	8.4% (991)	12.6% (1,495)	8.8% (1,042)	11.2% (1,333)
African	10.5% (298)	9.3% (265)	11.2% (320)	7.9% (224)	10.4% (297)
Latin American	7.4% (171)	9.2% (213)	8.3% (192)	4.4% (101)	5.7% (133)
South Central Asian	7.2% (1,396)	5.5% (1,078)	8.2% (1,587)	4.4% (863)	9.4% (1,835)
Northeast Asian	4.9% (928)	5.2% (978)	6.2% (1,166)	2.5% (471)	8.9% (1,671)
Southeast Asian	7.3% (845)	6.1% (711)	7.9% (924)	5.3% (614)	11.1% (1,292)
European	8.8% (12,248)	8.2% (11,415)	8.7% (12,147)	5.1% (7,032)	5.2% (7,272)
Child LBOTE					
No	9.5% (20,027)	8.5% (17,880)	9.2% (19,424)	5.8% (12,219)	5.9% (12,347)
Yes	9.2% (5,789)	7.4% (4,646)	10.4% (6,546)	7.7% (4,846)	13.1% (8,278)



Characteristics	PHY-WELL %(n)	EMO %(n)	SOC %(n)	LANG-COG %(n)	COM-KNO %(n)
Parent LBOTE					
No	9.1% (17,100)	8.3% (15,573)	8.9% (16,728)	5.4% (10,172)	5.6% (10,616)
Yes	8.0% (4,764)	6.7% (3,963)	9.3% (5,547)	6.2% (3,676)	11.5% (6,844)
Child not proficient in English					
No	7.9% (20,775)	7.3% (19,098)	8.0% (20,996)	4.6% (12,134)	4.0% (10,630)
Yes	46.3% (5,004)	31.7% (3,402)	45.8% (4,941)	45.6% (4,909)	≥ 90.0% (≥ 9,717) ^b

^a This group includes Anglo-Celtic and Australian backgrounds; ^b Data on developmental vulnerability on COM-KNO were suppressed due to 90 per cent vulnerability rule as per the AEDC Data Guidelines section 7.2.2.

Abbreviations: AEDC, Australian Early Development Census; COM-KNO Communication skills and general knowledge; EMO Emotional maturity; LANG-COG Language and cognitive skills (school-based); LBOTE Language background other than English; PHY-WELL Physical health and wellbeing; SOC Social competence.



Appendix B.

Part 2: Child-level and area-level disadvantage indicators

Table B.1. Distribution of DV1 by area-level child disadvantage (SEIFA)

Community SEIFA Quintiles	DV1 n (%)	
	No	Yes
Q5 (least disadvantaged)	30073 (87.3%)	4383 (12.7%)
Q4	36660 (85.1%)	6435 (14.9%)
Q3	41749 (83.1%)	8504 (16.9%)
Q2	46836 (79.5%)	12050 (20.5%)
Q1 (most disadvantaged)	58067 (69.8%)	25115 (30.2%)

Abbreviations: DV1 developmental vulnerability on one or more domain(s); Q quintiles.

Table B.2. Associations of area-level and child-level disadvantage indicators with children's DV1.

Indicators		DV1
		RR (95% CI)
Area-level		
SEIFA:		
	Q5 (least disadvantaged)	Reference
	Q4	1.17 (1.13, 1.22)
	Q3	1.33 (1.29, 1.38)
	Q2	1.61 (1.56, 1.66)
	Q1 (most disadvantaged)	2.37 (2.30, 2.45)
Child-level		
Sociodemographic		
Poverty line:		
	Above poverty line	Reference
	Poverty line or below	1.26 (1.24, 1.29)
Family eligible for a Low Income Health Care Card		
1.77 (1.74, 1.80)		
Family Tax Benefit A, based on income group:		
	Greater than \$99,864	Reference
	\$56,137 to \$99,864	1.58 (1.55, 1.62)
	\$56,137 or less	1.99 (1.95, 2.02)
Family Tax Benefit B, based on income group:		
	Greater than \$100,900	Reference
	\$100,900 or less	1.81 (1.78, 1.83)



Indicators		DV1
		RR (95% CI)
Child Care Subsidy, based on income group:		
	Greater than \$254,305	Reference
	Greater than \$175,015 to \$254,305	1.19 (1.13, 1.25)
	Greater than \$70,015 to \$175,015	1.62 (1.55, 1.69)
	\$0 to \$70,015	2.60 (2.49, 2.71)
Child Care Subsidy with income threshold:		
	Greater than \$70,015	Reference
	\$70,015 or less	1.77 (1.74, 1.79)
Parental highest education level:		
	Bachelor's degree or above	Reference
	Advanced Diploma or Diploma	1.47 (1.44, 1.50)
	Certificate level I to IV ^a	1.72 (1.68, 1.75)
	Year 12 or below	2.37 (2.33, 2.42)
Parent highest occupation:		
	Managers / Professionals	Reference
	Technicians / Other types of workers ^b	1.44 (1.41, 1.47)
	Labourers / Others ^c	1.66 (1.62, 1.70)
Parent highest occupation:		
	White collar	Reference
	Blue collar	1.37 (1.34, 1.40)
Parent was not employed		
		1.56 (1.51, 1.60)
Parent employment average duration:		
	Greater than 4 years	Reference
	4 years or less	1.72 (1.70, 1.75)
Parent received social support payment:		
	Age pension support payment	1.81 (1.61, 2.03)
	Carer support payment	1.65 (1.61, 1.68)
	Rent assistance support payment	1.76 (1.74, 1.79)
	Family support payment	1.54 (1.48, 1.59)
	Employment support payment	1.80 (1.77, 1.83)
	Student support payment	1.39 (1.33, 1.45)
	Disability support payment	1.91 (1.86, 1.97)
	Any type of social security payments	1.57 (1.51, 1.63)
Parent received special childcare benefit:		
	At risk childcare benefit	2.16 (2.10, 2.23)
	Financial hardship childcare benefit	1.63 (1.58, 1.68)
	Grandparent childcare benefit	2.34 (2.2, 2.49)
	Jobs education and training childcare benefit	1.59 (1.54, 1.64)
	Any special childcare benefit payments	1.79 (1.76, 1.83)
Child with a lone parent family		
		1.61 (1.59, 1.64)
Household size with 6 or more people		
		1.02 (0.99, 1.06)



Indicators		DV1
		RR (95% CI)
Health conditions		
Parent has had any chronic health issue(s)		1.23 (1.21, 1.25)
Child has had any chronic health issue(s)		1.59 (1.56, 1.62)
Parent has had any mental health issue(s)(MHI)		1.22 (1.20, 1.24)
Parent mental health issue duration:		
	No MHI issues or had MHI for less than one year	Reference
	Greater than one year	1.24 (1.23, 1.26)
Child has had any mental health issue(s)		1.82 (1.78, 1.86)
Child mental health issue duration:		
	No MHI issues or had MHI for less than one year	Reference
	Greater than one year	1.82 (1.75, 1.88)
Geographic		
House crowding (3 or more additional bedrooms needed)		2.01 (1.87, 2.16)
House crowding (1 or more additional bedrooms needed)		1.58 (1.55, 1.62)
Dwelling type:		
	Private dwellings	Reference
	Collective dwellings	0.97 (0.83, 1.14)
Tenure type:		
	Own	Reference
	Rented ^d	1.64 (1.62, 1.67)
Child has moved residence address in the last 5 years		1.14 (1.12, 1.16)
Risk factors		
Preschool non-attendance		1.79 (1.75, 1.82)
Childcare non-attendance		1.03 (1.02, 1.05)
Unpaid childcare		0.62 (0.61, 0.64)
Child's age group at childcare entry:		
	0-2 years	Reference
	3-6 years	1.10 (1.08, 1.12)
Child is not regularly read to at home		4.06 (4.00, 4.11)
Mother's age is 20 years or younger		1.92 (1.87, 1.97)
Mother's age is 35 years or older		0.84 (0.83, 0.86)
Parental death		1.59 (1.42, 1.78)

^a Certificate level I to IV including trade qualification. ^b Technicians, trade workers, community and personal service workers, and sales workers. ^c Clerical and administrative workers, machinery operators and driver, and labourers. ^d The tenure category 'living in residences that were rented' also includes 'occupied rent free', 'occupied under a life tenure scheme', and 'other tenure type'. *Note.* The results for child-level indicators were obtained in Phase One and are included here for comparison to the area-level indicator for completeness. *Abbreviations:* AEDC, Australian Early Development Census; CI Confidence Interval; DV1 Developmental vulnerability on one or more domain(s); MHI Mental health issues; Q quintiles; RR Risk ratio; SEIFA Socio-economic indexes for areas.



Appendix C.

Part 3: Best child-level disadvantage predictors of DV₁

Table C.1. Summary of child-level disadvantage indicators in the sociodemographic lens as predictors of DV₁ ordered by magnitude of univariable association.

Indicators	Description	Categories	DV ₁ RR (95% CI)	% Missing
Child Care Subsidy income thresholds	Child Care Subsidy was defined based on the raw household income.	0: \$254,305 or more;	Reference	10.3%
		1: \$175,015 to \$254,305;	1.19 (1.13, 1.25)	
		2: \$70,015 to \$175,015;	1.62 (1.55, 1.69)	
		3: \$70,015 or less	2.60 (2.49, 2.71)	
Parental highest education level	Parental highest educational level (either Parent 1 or Parent 2) in the household.	0: Bachelor's degree or above	Reference	11.6%
		1: Advanced Diploma or Diploma	1.47 (1.44, 1.50)	
		2: Certificate level I to IV (including trade qualification);	1.72 (1.68, 1.75)	
		3: Year 12 or below	2.37 (2.33, 2.42)	
Family Tax Benefit A	Family Tax Benefit A was defined based on the raw household income.	0: \$99,864 or more;	Reference	10.3%
		1: \$56,137 to \$99,864;	1.58 (1.55, 1.62)	
		2 \$56,137 or less	1.99 (1.95, 2.02)	
Family Tax Benefit B	Family Tax Benefit B was defined based on the raw household income.	0: >\$100,000;	Reference	10.3%
		1: <=\$100,000	1.81 (1.78, 1.83)	
Parent received special childcare benefit	Family received any special childcare benefit, which includes at-risk childcare benefit, financial hardship childcare benefit, grandparent childcare benefit, or jobs education and training childcare fee assistance.	0: No;	Reference	30.1%
		1: Yes	1.79 (1.76, 1.83)	
Family eligible for Low Income Health care card	Low Income Health Care Card was defined based on the raw household income and household composition (Single-parent household: <=\$71,955; Two-parent household: <= \$74,165).	No	Reference	10.5%
		Yes	1.77 (1.74, 1.80)	



Indicators	Description	Categories	DV1 RR (95% CI)	% Missing
Parent employment average duration	Parental employment duration (either Parent 1 or Parent 2) based on an average four year cut off.	0: >4 years;	Reference	8.5%
		1: =<4 years	1.72 (1.70, 1.75)	
Parent highest occupation	Parental highest occupation (either Parent 1 or Parent 2) with three categorises in the household.	0: Managers/Professionals;	Reference	30.3%
		1: Technicians/Trade owners/ Community and Personal Service workers/Sale workers;	1.44 (1.41, 1.47)	
		2: Clerical and Administrative workers/ Machinery operators/Drivers and labourers	1.66 (1.62, 1.70)	
Child with a lone parent family	The household has a single parent.	0: No;	Reference	9.2%
		1: Yes	1.61 (1.59, 1.64)	
Parent received social support payment	Family received any type of social support payment, which includes age pension, carer payment, rent assistance, family support (i.e., it does not include Baby Bonus, Child Care Benefits, family allowance or maternity payments, but does include family tax benefits), unemployment payment, student support, or disability support.	0: No;	Reference	6.9%
		1: Yes	1.57 (1.51, 1.63)	
Parent employment status	Parental employment status (either Parent 1 or Parent 2) in the household.	0: Employed;	Reference	26.5%
		1: Not employed	1.56 (1.51, 1.60)	
Parent highest occupation	Parental highest occupation (either Parent 1 or Parent 2) with two categorises in the household.	0: White collar (Managers, Professionals, Community and personal service workers, clerical and administrative workers, and sale workers);	Reference	30.3%
		1: Blue collar (Technicians/ trade workers, machinery operators, drivers, and labourers)	1.37 (1.34, 1.40)	
Poverty line	Poverty line defined as 50% or less of the median equivalised household income (i.e., \$41,092.64).	0: Above poverty line;	Reference	10.4%
		1: At or below poverty line	1.26 (1.24, 1.29)	
Household size with 6 or more people	Number of people living in the household with a 5-person cut off.	0: 5 people or less;	Reference	10.1%
		1: 6 people or more	1.02 (0.99, 1.06)	

Abbreviation: CI Confidence Interval; DV1 developmental vulnerability on one or more domain(s); RR Risk ratio.



Correlation plot

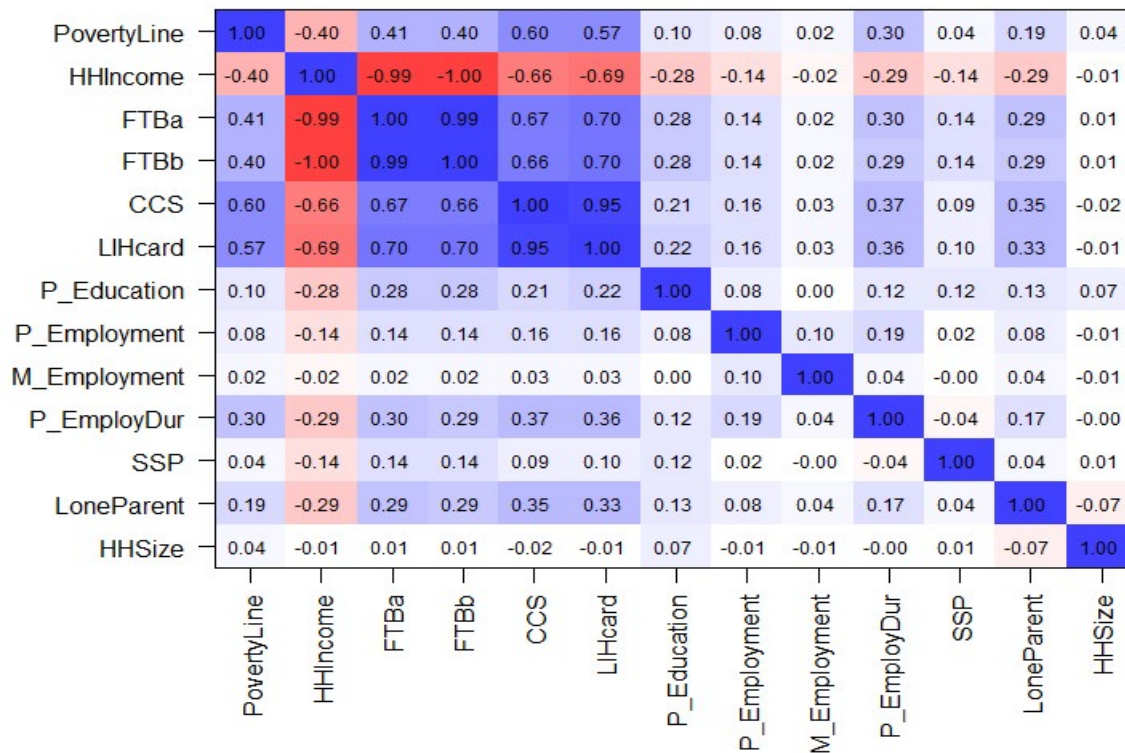


Figure C.1. Correlation matrix for child-level disadvantage indicators in the sociodemographic lens.

Red indicates negative correlations and blue indicates positive correlations. *r* values indicate correlation strength: 0-0.39=weak, 0.40-0.59=moderate, 0.60-1.0=strong. *Abbreviations:* CCS Child Care Subsidy income thresholds; EmployDur Employment Duration; FTBa Family Tax Benefit A; FTBb Family Tax Benefit B; HHIncome Household income; LIHcard Low Income Health Care Card; M Maternal; P Parent; SSP Social support payment.

**Table C.2.** Summary of child disadvantage indicators in the health conditions lens as predictors of DV1 ordered by magnitude of univariable association.

Indicators	Variable description	Categories	DV1 RR (95% CI)	% Missing of total N=293,910
Child has had any mental health issue(s)	Child used mental health service/script access between birth and 2018.	0: No	Reference	7.0%
		1: Yes	1.82 (1.78, 1.86)	
Child mental health issue duration	The duration of child mental health issue between birth to 2018 (MBS/PBS).	0: Child has no mental health issues or has had a mental health issue less than one year	Reference	7.0%
		1: Child has had a mental health issue greater than one year	1.82 (1.75, 1.88)	
Child has had any chronic health issue(s)	Child used chronic health service/script access between birth and 2018.	0: No	Reference	7.0%
		1: Yes	1.59 (1.56, 1.62)	
Parent mental health issue duration	The duration of at least one parent with mental health service/script access between one year prior to birth and 2018.	0: Parent has no mental health issues or has had a mental health issue less than one year;	Reference	7.1%
		1: Parent has had a mental health issue greater than one year	1.24 (1.23, 1.26)	
Parent has had any chronic health issue(s)	At least one parent with chronic health service/script access between birth and 2018.	0: No	Reference	7.1%
		1: Yes	1.23 (1.21, 1.25)	
Parent has had any mental health issue(s)	At least one parent with mental health service/script access between one year prior to birth and 2018.	0: No	Reference	7.1%
		1: Yes	1.22 (1.20, 1.24)	

Abbreviations: CI Confidence Interval; DV1 developmental vulnerability on one or more domain(s); RR Risk ratio.

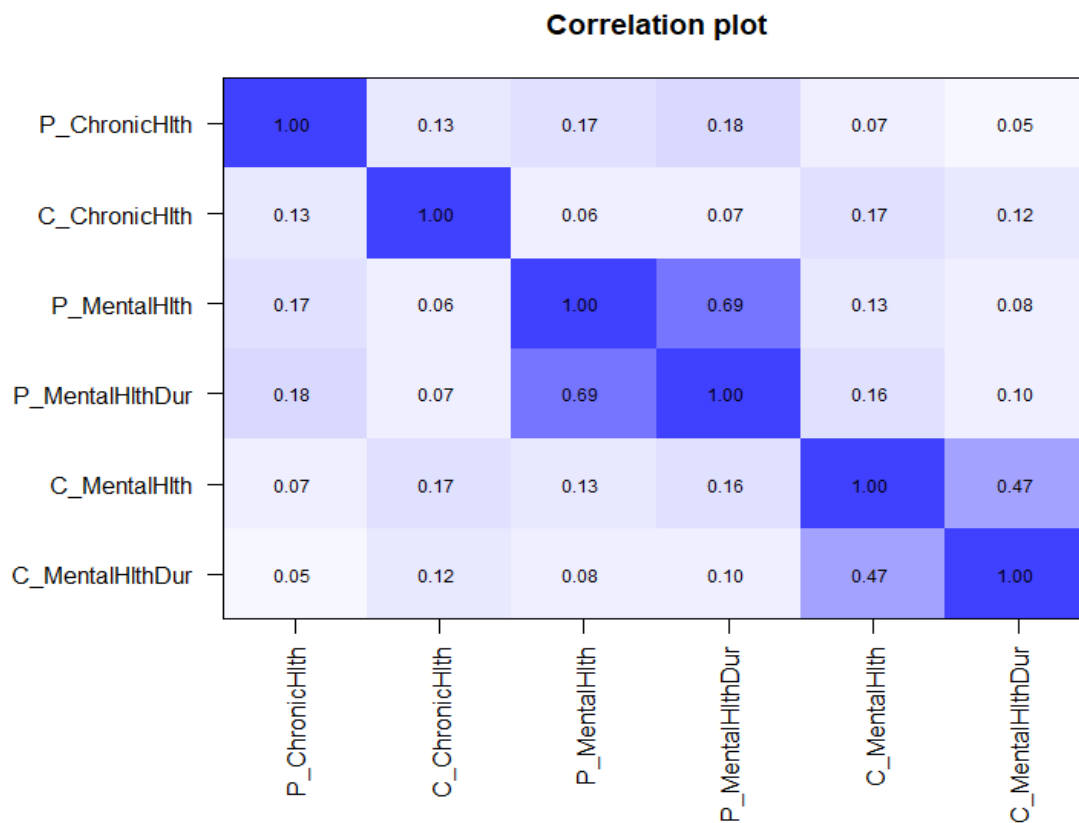


Figure C.2. Correlation matrix for child-level disadvantage indicators in the health conditions lens.

Blue indicates positive correlations. r values indicate correlation strength: 0-0.39=weak, 0.40-0.59=moderate, 0.60-1.0=strong. *Abbreviations:* C Child; MentalHlthDur Duration of mental health issues (in years); P Parent.



Table C.3. Summary of child disadvantage indicators in the geographic lens as predictors of DV1 ordered by magnitude of univariable association.

Indicators	Variable description	Categories	DV1 RR 95% CI	% Missing of total N=293,910
House crowding (3 or more additional bedrooms needed)	House crowding with three or more additional bedrooms needed in the household where the child lived.	0: 1-2 bedrooms extra bedrooms needed/ none needed/ spare bedrooms	Reference	23.6%
		1: 3 or more extra bedrooms needed	2.01 (1.87, 2.16)	
Tenure type	Tenure type was classified into the following categories: Owned outright,	0: Own (Owned outright, Owned with a mortgage, Being purchased under a shared equity scheme);	Reference	22.3%
	Owned with a mortgage, Being purchased under a shared equity scheme, Rented, Being occupied rent-free, Being occupied under a life tenure scheme, Other tenure type.	1: Rented (Rented, Being occupied rent-free, Being occupied under a life tenure scheme, Other tenure type)	1.64 (1.62, 1.67)	
House crowding (1 or more additional bedrooms needed)	House crowding with one or more additional bedrooms needed in the household where the child lived.	0: None needed/Spare bedrooms	Reference	23.6%
		1: One or more extra bedrooms needed	1.58 (1.55, 1.62)	
Child has moved residence address in the last 5 years	It indicates if all, some, or none of the usual residents of a household on 10 August 2021 have a different usual address compared to five years earlier (i.e., 10 August 2016).	0: No (No residents aged five years and over had a different address five years ago)	Reference	23.2%
		1: Yes (All residents in the household aged five years and over had a different address five years ago, or Some residents aged five years and over had a different address five years ago)	1.14 (1.12, 1.16)	
Dwelling type	Dwelling type was classified into the following categories: Occupied private dwellings, Non-private dwellings, Migratory, Off-shore, Shipping.	0: Collective dwellings (Non-private dwellings, Migratory, Off-shore, or	Reference	21.0%
		1: Occupied private dwellings	1.03 (0.88, 1.20)	

Abbreviation: CI Confidence Interval; DV1 developmental vulnerability on one or more domain(s); RR Risk ratio.

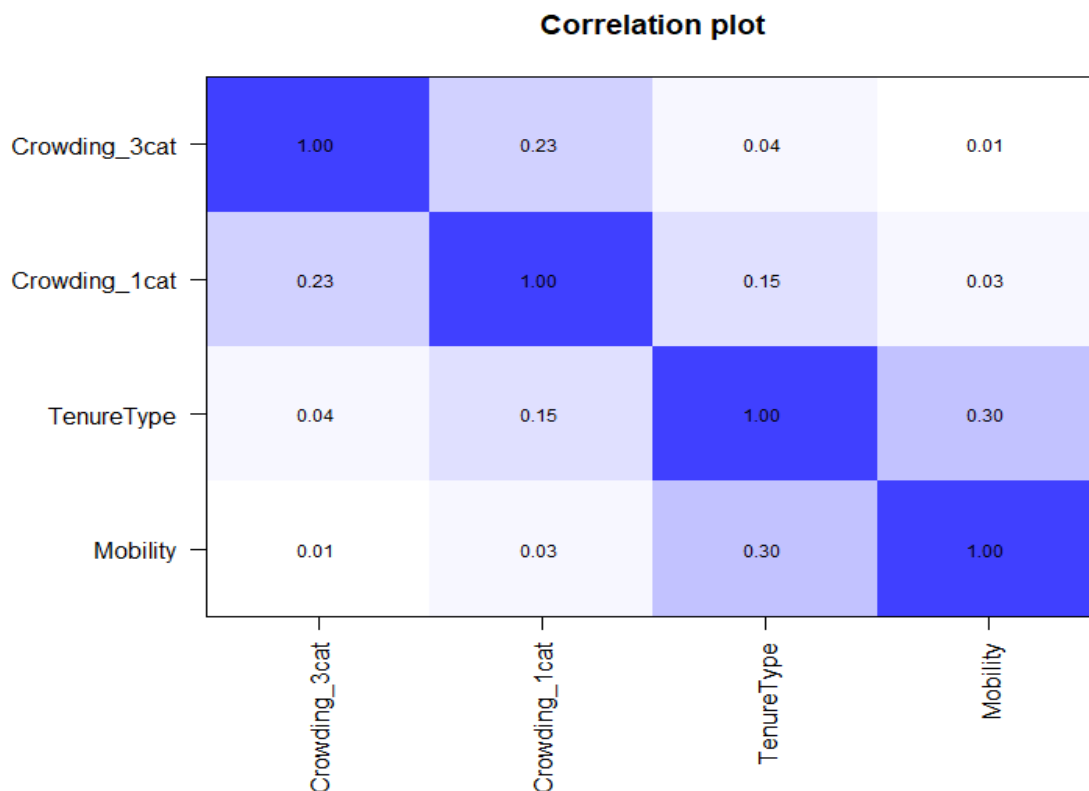


Figure C.3. Correlation matrix for child-level disadvantage indicators in the geographic lens.

Red indicates negative correlations and blue indicates positive correlations. R values indicate correlation strength: 0-0.39=weak, 0.40-0.59=moderate, 0.60-1.0=strong. *Abbreviations:* Cat categories.



Table C.4. Summary of child disadvantage indicators in the risk factor lens as predictors of DV1 ordered by magnitude of univariable association.

Indicators	Variable description	Categories	DV1 RR (95% CI)	% Missing of total N=293,910
Child home education environment	Child is regularly read to at home.	1: Yes (Very true or somewhat true)	Reference	9.5%
		0: No (Not Sure)	4.06 (4.00, 4.11)	
Mother's age at birth (teenager)	Maternal age at birth was identified using the month and year of birth for both the mother and the child. Two categories were created using the 20 years or younger as cut off.	0: >=20 years	Reference	8.2%
		1: <20 years	1.92 (1.87, 1.97)	
Preschool attendance	A child was defined as having attended preschool if they were marked as having attended preschool in the AEDC dataset and/or had at least 600 hours of Long Day Care (LDC) in the CCMS in the year before school.	1: Yes	Reference	10.9%
		0: No	1.79 (1.75, 1.82)	
Unpaid childcare	A child was identified as being exposed to unpaid childcare if any parent self-reported providing unpaid childcare for their own or other children in the past two weeks in the 2016 Census	0: Yes	Reference	16.2%
		1: No	1.61 (1.56, 1.64)	
Parental death	Parental death was identified as the death of a child's parent in the time period after the child's birth but prior to the completion of the AEDC.	0: No	Reference	6.9%
		1: Yes	1.59 (1.42, 1.78)	
Mother's age at birth (later)	Maternal age at birth was identified using the month and year of birth for both the mother and the child. Two categories were created using the 35 years or older as cut off.	0: >=35 years	Reference	8.2%
		1: <35 years	1.19 (1.16, 1.20)	
Child's age group at childcare entry	Child's age at childcare entry years was identified as the time difference between the first quarter of CCMS attendance and the child's birth.	0: 0-2 years	Reference	30.1%
		1: 3 or more years	1.10 (1.08, 1.12)	
Childcare attendance	Childcare attendance was identified using data from the CCMS. If a child had a record in the CCMS at any time, they were flagged as having attended childcare.	1: Yes	Reference	6.9%
		0: No	1.03 (1.02, 1.05)	

Abbreviation: CI Confidence Interval; DV1 developmental vulnerability on or more one domain(s); RR Risk ratio.



Correlation plot

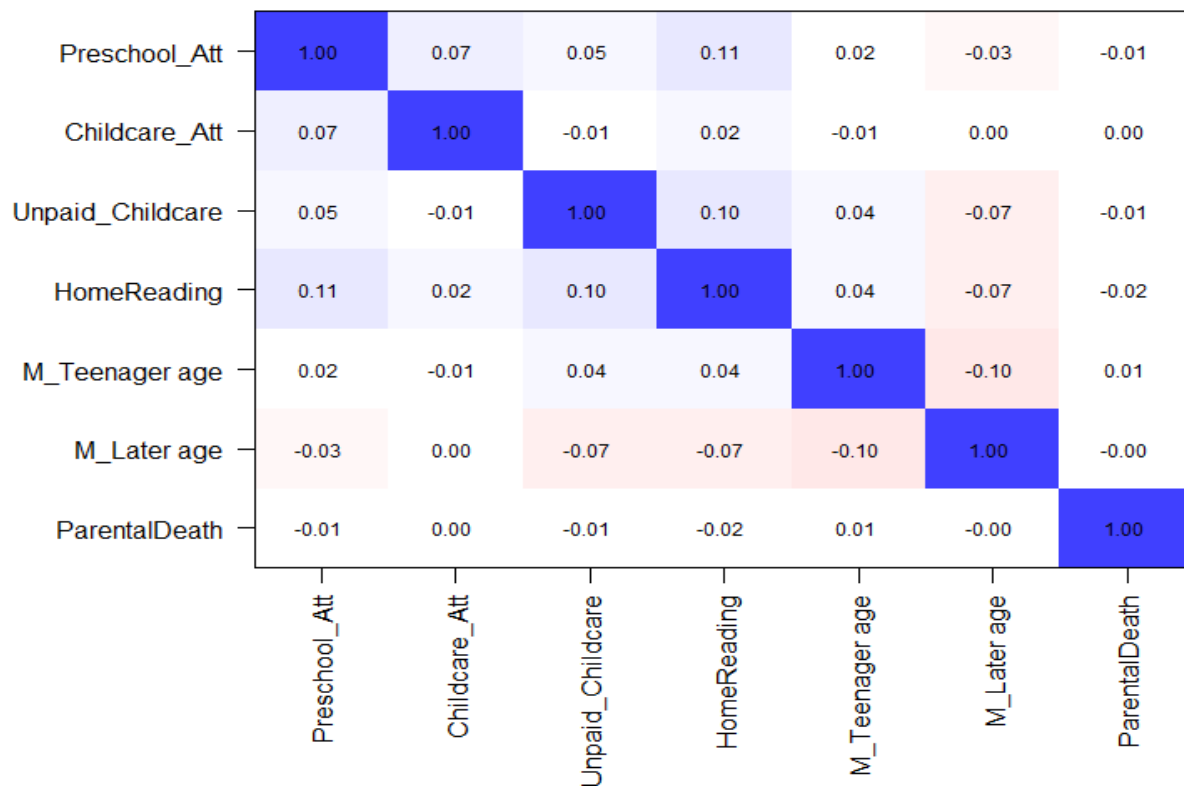


Figure C.4. Correlation matrix for child disadvantage indicators in the risk factors lens

Red indicates negative correlations and blue indicates positive correlations. R values indicate correlation strength: 0-0.39=weak, 0.40-0.59=moderate, 0.60-1.0=strong. *Abbreviations:* Att Attendance; M Maternal.

**Table C.5. Summary of priority population indicators as predictors of DV1 ordered by magnitude of univariable association.**

Indicators	Variable description	Categories	DV1 RR (95% CI)	% Missing of total N=293,910
Child not proficient in English	Child is not proficient in English	0: No (Very good, good, average)	Reference	0.1%
		1: Yes (Very poor or poor)	5.29 (5.24, 5.34)	
Parent ancestry	Parent ancestry with 10 categories (Census) 2016 (prioritisation)	0: Anglo-Celtic ^a	Reference	16.8%
		1: Aboriginal and/or Torres Strait Islander	2.03 (1.93, 2.13)	
		2: Pacific Islander/Māori	1.26 (1.20, 1.32)	
		3: Middle Eastern	1.19 (1.15, 1.24)	
		4: African	1.15 (1.08, 1.23)	
		5: Latin American	0.86 (0.79, 0.94)	
		6: South Central Asian	0.87 (0.84, 0.90)	
		7: Northeast Asian	0.76 (0.73, 0.79)	
		8: Southeast Asian	0.94 (0.90, 0.98)	
		9: European	0.88 (0.86, 0.90)	
Child's LBOTE	Child has a language background other than English.	0: No	Reference	6.9%
		1: Yes	1.20 (1.18, 1.22)	
Child's country of birth	Country of child's birth was identified using the country of birth indicator available in the combined demographics file. Country of child's birth was categorised into three groups: Australia, other Organisation for Economic Cooperation and Development (OECD), and non-OECD.	0: 0: Australia	Reference	7.0%
		1: Other English-speaking country (Other OECD)	0.98 (0.92, 1.04)	
		2: Other non-English speaking country (non-OECD)	1.19 (1.14, 1.23)	
Parents LBOTE	At least one parent has a language background other than English	0: No	Reference	15.8%
		1: Yes	1.13 (1.11, 1.15)	

^a This group includes Anglo-Celtic and Australian backgrounds. *Abbreviations:* CI Confidence Interval; DV1 developmental vulnerability on one or more domain(s); LBOTE Language background other than English; OECD Organisation for Economic Cooperation and Development; RR Risk ratio.

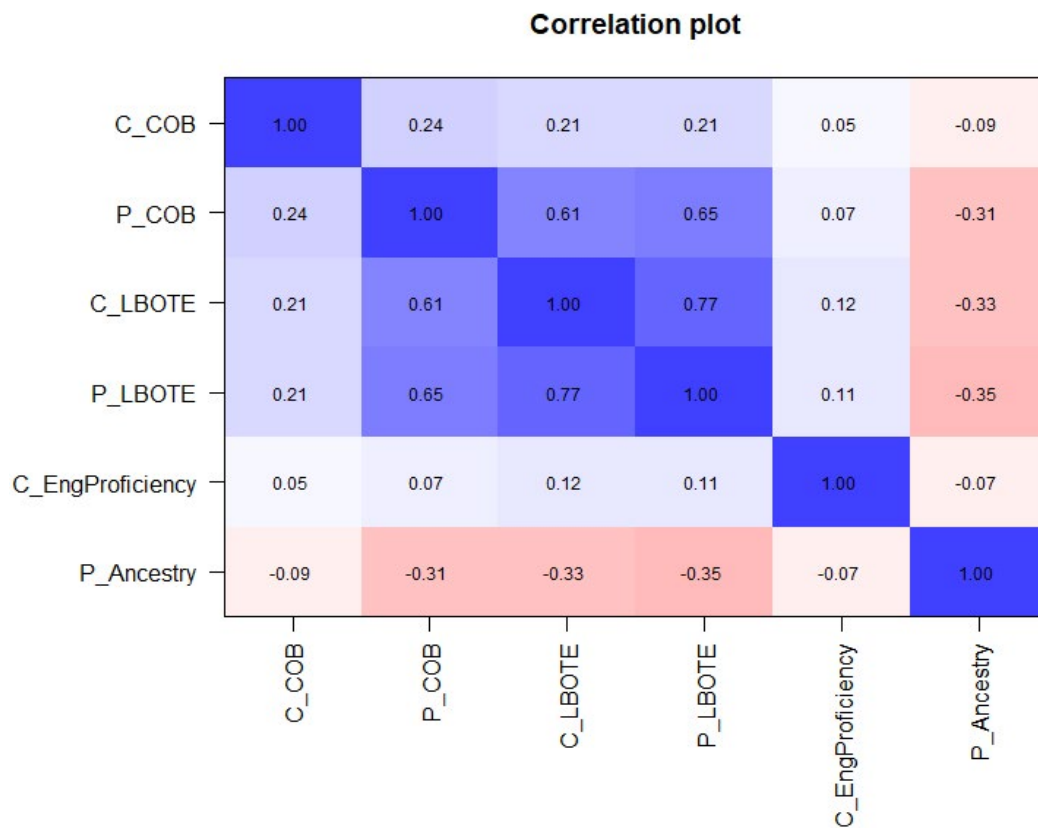


Figure C.5. Correlation matrix for priority population indicators.

Red indicates negative correlations and blue indicates positive correlations. R values indicate correlation strength: 0-0.39=weak, 0.40-0.59=moderate, 0.60-1.0=strong. *Abbreviations:* C Child; COB Country of birth; LBOTE Language background other than English; P Parent.



The Centre for Community Child Health
The Royal Children's Hospital Melbourne
50 Flemington Road
Parkville 3052 VIC

www.ccch.org.au

The Centre for Community Child Health is a department of The Royal Children's Hospital
and a research group of the Murdoch Children's Research Institute.