



Waipapa  
Taumata Rau  
University  
of Auckland



# Growing Up in New Zealand Data in the Statistics New Zealand Integrated Data Infrastructure

A technical document for data users

v1.2 2026



**Making Aotearoa a better place to call home**

# Contents

|                                                                                                       |   |
|-------------------------------------------------------------------------------------------------------|---|
| 1. Acronyms and abbreviations . . . . .                                                               | 4 |
| 2. Purpose and scope of this report . . . . .                                                         | 5 |
| 3. Intended users of this report . . . . .                                                            | 5 |
| 3.1 Recommended reading and supporting resources . . . . .                                            | 5 |
| 3.2 General Considerations . . . . .                                                                  | 5 |
| 4. GUiNZ IDI sub-set . . . . .                                                                        | 6 |
| 4.1 Participant Consent . . . . .                                                                     | 6 |
| 4.2 Characteristics of the GUiNZ IDI-consented cohort sample compared with the total cohort . . . . . | 6 |
| 4.3 Characteristics of the GUiNZ IDI consented cohort, by ethnicity . . . . .                         | 8 |
| 5. Summary . . . . .                                                                                  | 9 |
| 6. Useful References . . . . .                                                                        | 9 |

# Tables

Table 1. Characteristics of the GUiNZ (young people) IDI-consented sample and wider cohort . . . . . 7

Table 2. Characteristics of the GUiNZ (young people) IDI-consented sample by ethnicity\* . . . . . 8

# 1. Acronyms and abbreviations

|           |                                                 |
|-----------|-------------------------------------------------|
| DCW       | Data Collection Wave                            |
| GUiNZ     | Growing Up in New Zealand                       |
| IDI       | Integrated Data Infrastructure                  |
| MELAA     | Middle Eastern, Latin American and African      |
| NCEA      | National Certificate of Educational Achievement |
| NHI       | National Health Index                           |
| NZDep2006 | New Zealand Deprivation Index 2006              |
| PC        | Primary caregiver                               |
| SNZ       | Statistics New Zealand                          |
| UoA       | University of Auckland                          |

## 2. Purpose and scope of this report

A linked sub-set of Growing Up in New Zealand (GUiNZ) study data are now available in the Statistics New Zealand (SNZ) Integrated Data Infrastructure (IDI). Linking the GUiNZ sub-set of data with other de-identified data sources in the IDI creates opportunities to better understand the individual, family, community, service, and wider societal factors that support the wellbeing and development of children and young people in New Zealand. In this way, the availability of GUiNZ data in the IDI supports the study's overarching aims by enabling researchers to examine children's and young people's experiences and outcomes across multiple areas of their lives over time.

**The appropriate use of GUiNZ data in the IDI requires a clear understanding of who is included in the sub-set, who is excluded, and whether the sub-set sample is suitable for your research question. This document sets out some of these considerations.**

Importantly, data users should understand whether the sub-set of GUiNZ participants included in the IDI dataset differ systematically from those in the full GUiNZ dataset. Differences in consent, linkage, and ongoing cohort participation may affect the representativeness of the IDI sample and, therefore, the interpretation of analyses based on linked data. The extent of this impact will depend on the research questions and analytical approaches used.

The purpose of this report is to provide IDI data users with information on participants included in the GUiNZ consented sub-set to help assess the suitability of the data for answering research questions and identifying any analytic adjustments that may be required. This is useful to inform the planning and interpretation of IDI-based analysis and findings.

## 3. Intended users of this report

This report is intended for anyone who plans to use the GUiNZ data within the SNZ IDI. It provides information on the sub-set who consented for data to be made available in the IDI, including comparison to the total GUiNZ cohort and the potential implications of any differences for analysis and interpretation. For those planning to use the full GUiNZ data sets (outside the IDI), the technical information about the data is available on the [GUiNZ website](#), and this information will also be useful for those working with GUiNZ data in the IDI environment.

### 3.1 Recommended reading and supporting resources

Before planning analyses using the GUiNZ data available in the IDI, data users should familiarise themselves with the original GUiNZ cohort profile<sup>1</sup> and cohort-alignment paper<sup>2</sup>, including the reports and questionnaires for relevant data collection waves, the associated data dictionaries, and the GUiNZ Data User Guide<sup>3</sup>. These resources provide important context about the original cohort design, recruitment population, data collection methods, available variables, and longitudinal participation patterns<sup>4</sup>. They should be read alongside this report when assessing whether the GUiNZ IDI-consented sub-set is suitable for a proposed research question. Relevant documents are listed in the Useful References section.

The GUiNZ sampling strategy was designed so that evidence provided over time would be broadly generalisable to the population of New Zealand births at the time of recruitment.

### 3.2 General Considerations

Tables 1 and 2 provide descriptive information to help data users assess the composition of the GUiNZ IDI-consented sub-set and its suitability to answer a proposed research question. The IDI-consented sub-set should not be assumed to be a random or fully representative subset of the original GUiNZ cohort. Differences in IDI consent, ongoing eligibility and cohort participation at each data collection wave (DCW), and the completeness of data across collection waves may affect which participants can be included in any particular analysis, and how findings should be interpreted.

Data users should, before undertaking their analysis, clearly define the population to which their research question relates and identify the GUiNZ longitudinal measures and IDI data required. Data users should then assess whether the required data, such as ethnic groups, collection waves, and subgroup data, are adequately represented in the available analytical sample.

The descriptive differences presented in this document do not, on their own, establish the presence of selection bias. Further analysis is required to understand the factors contributing to differential consent and inclusion in the IDI consented sub-cohort. Until that work is completed, findings should be interpreted as applying primarily to the participants included in the relevant analytical sample and should not be generalised to the same target population as analyses that could use the full GUiNZ cohort without further consideration.

Depending on the research question and the observed composition of the analytical sample, data users may need to consider approaches such as stratified analyses, covariate adjustment, or sensitivity



analyses. The choice of approach should be guided by the research question and the available sample size.

Data users will discover that some subgroups within the IDI-consented sample are small, and will result in reduced statistical power and precision of estimates, and the increased likelihood of output suppression under SNZ confidentiality requirements.

For analyses examining outcomes according to ethnicity, particular attention should be given to subgroup sample sizes, the method used to classify ethnicity, and whether estimates are sufficiently precise to support ethnicity-specific interpretation. Observed differences by ethnicity should be described and interpreted cautiously.

## 4. GUiNZ IDI sub-set

### 4.1 Participant Consent

The GUiNZ baseline cohort includes 6,853 children who were born to 6,823 mothers between March 2009 and April 2010 from a region spanning Auckland, Counties Manukau and Waikato District Health Boards (DHBs).<sup>5</sup> Although consent for bilateral linkage of the GUiNZ longitudinal data to administrative datasets at the University of Auckland has been sought and updated over time, the potential for linkage and analysis of the GUiNZ dataset in the IDI had not previously been obtained, including at recruitment as the IDI did not exist at that time. To facilitate data sharing of the entire existing longitudinal dataset in the IDI, a consent process was designed and undertaken by the GUiNZ team in 2025. Full details of this consent process will be provided in an upcoming paper but is outlined briefly below.

Consent to IDI linkage was sought from all eligible GUiNZ participants. The NWA15 Methods paper contains the definition of the eligible GUiNZ cohort.<sup>6</sup>

Consent for sharing and IDI linkage of all existing GUiNZ data (collected between pregnancy up to and including age 15-years) was sought from both the primary caregiver and their child. Primary caregiver consent included consent for sharing of their own personal health and wellbeing data, household level data and health and wellbeing data for their child. Child consent included consent for sharing their own personal health and wellbeing data. Given the age

of the child participants at the time of this consent process, and the fact that child-level data had initially been provided by caregivers as well as more recently being provided directly by children (primary caregivers from the AN wave to 15Y; children from 8Y to 15Y), it was determined that consent from both respondent types would be required for child data to be shared. A total of 1493 GUiNZ caregivers and young person dyads submitted a consent form agreeing for data linkage (i.e. where both the primary caregiver and the young person consented).

The GUiNZ IDI consent process (carried out between July and October 2025) represents the first stage in making GUiNZ data available for linkage within the SNZ IDI. There will be a future update to the GUiNZ data within the SNZ IDI following the completion of the 17Y DCW, and we expect that any further DCWs will seek consent for new data to be shared to enable linked analyses. Data users should confirm with SNZ which GUiNZ DCWs and linked variables are available in the IDI at the time of developing their projects.

### 4.2 Characteristics of the GUiNZ IDI-consented cohort sample compared with the total cohort

Table 1 presents the baseline demographic, socioeconomic, and geographic characteristics of the young people who consented to IDI linkage (the “IDI consented” cohort) and compares this group with the GUiNZ cohort who participated in three previous data collection waves: the Antenatal DCW (2009-10)<sup>5</sup>, 12-year DCW (2021-22)<sup>7</sup> and the 15-year DCW (2025)<sup>8</sup>. The table is intended to help data users assess the differences in baseline characteristics between the IDI-consented group and the wider GUiNZ cohort.

Table 1 shows that the IDI-consented sample should not be considered a random subset of the antenatal cohort. Compared to the antenatal cohort, there are differences across the DCWs in ethnicity, age, higher educational qualifications, and deprivation quintile.

Specifically, the IDI-consented sample has higher representation for European participants, older mothers, mothers with higher educational qualifications, and participants living in the less deprived areas (quintiles 1 and 2) compared to the antenatal cohort. Data users should therefore consider the effects of these differences on their research question.

**Table 1. Characteristics of the GUiNZ (young people) IDI-consented sample and wider cohort**

| Characteristic                                 | Antenatal<br>(N=6822) | 12Y DCW<br>(N=4787) | 15Y DCW<br>(N=4183) | IDI Consented*<br>(N=1493) |
|------------------------------------------------|-----------------------|---------------------|---------------------|----------------------------|
| <b>Child sex at birth</b>                      |                       |                     |                     |                            |
| Male                                           | 3562 (52.0%)          | 2459 (51.4%)        | 2,128 (51.0%)       | 765 (51.0%)                |
| Female                                         | 3320 (48.0%)          | 2328 (48.6%)        | 2,020 (49.0%)       | 728 (49.0%)                |
| <b>Ethnicity (mother)#</b>                     |                       |                     |                     |                            |
| Māori                                          | 1244 (18.3%)          | 767 (16.0%)         | 674 (16.0%)         | 185 (12.3%)                |
| Pacific                                        | 1155 (17.0%)          | 485 (10.1%)         | 427 (10.0%)         | 94 (6.3%)                  |
| Asian                                          | 1081 (15.9%)          | 640 (13.4%)         | 606 (15.0%)         | 189 (12.7%)                |
| European                                       | 4242 (62.3%)          | 2,722 (56.9%)       | 2291 (56.0%)        | 971 (65.0%)                |
| Other                                          | 39 (0.6%)             | 167 (3.5%)          | 69 (1.7%)           | 27 (1.8%)                  |
| <b>Maternal age at antenatal DCW</b>           |                       |                     |                     |                            |
| <25 years                                      | 1325 (19.4%)          | 687 (14.4%)         | 578 (14.3%)         | 162 (10.9%)                |
| 25-29 years                                    | 1666 (24.2%)          | 1,127 (23.5%)       | 949 (23.0%)         | 325 (22%)                  |
| 30-34 years                                    | 2121 (31.1%)          | 1,656 (34.6%)       | 1444 (35.0%)        | 557 (37%)                  |
| >=35 years                                     | 1715 (25.0%)          | 1316 (27.5%)        | 1156 (28.0%)        | 447 (30.1%)                |
| <b>Maternal education at birth</b>             |                       |                     |                     |                            |
| None or secondary                              | 2118 (31.1%)          | 1212 (25.3%)        | 1048 (25.8%)        | 295 (19.8%)                |
| Diploma or certificate                         | 2083 (30.6%)          | 1410 (29.5%)        | 1169 (28.0%)        | 408 (27%)                  |
| Bachelor's or higher                           | 2603 (38.2%)          | 2151 (44.9%)        | 1901 (46.0%)        | 786 (53%)                  |
| <b>Area socioeconomic deprivation at birth</b> |                       |                     |                     |                            |
| Quintile 1 (least deprived)                    | 1092 (16.4%)          | 903 (18.9%)         | 763 (18.0%)         | 297 (20%)                  |
| Quintile 2                                     | 1158 (17.4%)          | 990 (20.7%)         | 834 (20.0%)         | 333 (22%)                  |
| Quintile 3                                     | 1160 (17.5%)          | 899 (18.8%)         | 769 (19.0%)         | 310 (21%)                  |
| Quintile 4                                     | 1405 (21.2%)          | 959 (20.0%)         | 852 (21.0%)         | 292 (20%)                  |
| Quintile 5 (most deprived)                     | 1828 (27.5%)          | 1034 (21.6%)        | 908 (22.0%)         | 259 (17%)                  |
| <b>Region at birth</b>                         |                       |                     |                     |                            |
| Auckland                                       | 2423 (35.5%)          | 2725 (56.9%)        | 1555 (38.0%)        | 616 (41%)                  |
| Counties Manukau                               | 2526 (37.0%)          | 1377 (27.8%)        | 1330 (32.0%)        | 411 (28%)                  |
| Waikato                                        | 1873 (27.5%)          | 1297 (27.1%)        | 1242 (30.0%)        | 465 (31%)                  |
| <b>Urban/rural at birth</b>                    |                       |                     |                     |                            |
| Urban                                          | 6378 (93.4%)          | 4,405 (92.0%)       | 3,803 (92.0%)       | 1,361 (91%)                |
| Rural                                          | 478 (7.0%)            | 381 (8.0%)          | 324 (7.9%)          | 131 (8.8%)                 |

\*Number of consented young people with NHI at antenatal (n=1493)

# Maternal self-reported ethnicity is used for consistency given children couldn't report ethnicity prior to 8Y

## 4.3 Characteristics of the GUiNZ IDI consented cohort, by ethnicity

Table 2 shows the characteristics of the young people who consented to IDI linkage, overall and by maternal ethnicity. This table is intended to help data users see the composition of the consented sample within the main ethnic groups represented in the study.

Table 2 provides descriptive information to support assessment of the suitability of the IDI-consented sample for ethnic-specific or subgroup analyses, particularly for broad comparisons by maternal age, maternal education, deprivation, region, and urban/rural location. Data users in the IDI should therefore assess subgroup sample sizes before analysis, as small numbers may limit statistical power and precision and increase the likelihood of output suppression under SNZ confidentiality rules.

**Table 2. Characteristics of the GUiNZ (young people) IDI-consented sample by ethnicity\***

| Characteristic                        | Māori<br>(N=185) | Pacific<br>(N=94) | Asian<br>(N=189) | European<br>(N=971) | Overall<br>(N=1493) |
|---------------------------------------|------------------|-------------------|------------------|---------------------|---------------------|
| <b>Child sex</b>                      |                  |                   |                  |                     |                     |
| Male                                  | 95 (51%)         | 51 (54%)          | 95 (50%)         | 510 (51.1%)         | 765 (51%)           |
| Female                                | 90 (49%)         | 43 (46%)          | 94 (50%)         | 488 (48.9%)         | 728 (49%)           |
| <b>Maternal age</b>                   |                  |                   |                  |                     |                     |
| <25 years                             | 48 (26.4%)       | 21 (22.1%)        | 17 (9.0%)        | 76 (7.6%)           | 162 (10.9%)         |
| 25-29 years                           | 47 (25%)         | 23 (24%)          | 62 (33%)         | 190 (19.0%)         | 325 (22%)           |
| 30-34 years                           | 46 (25%)         | 29 (31%)          | 73 (39%)         | 402 (40.3%)         | 557 (37%)           |
| >=35 years                            | 44 (24.2%)       | 21 (22.4%)        | 37 (19.6%)       | 330 (33.1%)         | 447 (30.1%)         |
| <b>Maternal education</b>             |                  |                   |                  |                     |                     |
| None or secondary                     | 64 (34.2%)       | 34 (36.3%)        | 32 (17.1%)       | 158 (15.8%)         | 295 (19.8%)         |
| Diploma or certificate                | 62 (34%)         | 32 (34%)          | 46 (24%)         | 263 (26.4%)         | 408 (27%)           |
| Bachelor's or higher                  | 59 (31.6%)       | 28 (29.5%)        | 111 (59.0%)      | 577 (58.4%)         | 786 (53%)           |
| <b>Material deprivation</b>           |                  |                   |                  |                     |                     |
| No                                    | 72 (39%)         | 30 (32%)          | 101 (53%)        | 521 (52.2%)         | 741 (50%)           |
| Yes                                   | 107 (57.7%)      | 61 (64.6%)        | 81 (42.8%)       | 471 (47.7%)         | 728 (48.6%)         |
| Missing                               | 6 (3.2%)         | 3 (3.2%)          | 7 (3.7%)         | 6 (0.6%)            | 24 (1.6%)           |
| <b>Area socioeconomic deprivation</b> |                  |                   |                  |                     |                     |
| Quintile 1 (least deprived)           | 21 (11%)         | 3 (3.2%)          | 22 (12%)         | 243 (24.3%)         | 297 (20%)           |
| Quintile 2                            | 29 (16%)         | 7 (7.4%)          | 31 (16%)         | 261 (26.2%)         | 333 (22%)           |
| Quintile 3                            | 37 (20%)         | 9 (9.6%)          | 37 (20%)         | 222 (22.2%)         | 310 (21%)           |
| Quintile 4                            | 34 (18%)         | 23 (24%)          | 63 (33%)         | 167 (16.7%)         | 292 (20%)           |
| Quintile 5 (most deprived)            | 64 (35%)         | 52 (55%)          | 36 (19%)         | 104 (10.4%)         | 259 (17%)           |
| <b>Region</b>                         |                  |                   |                  |                     |                     |
| Auckland                              | 36 (19%)         | 35 (37%)          | 98 (52%)         | 435 (43.6%)         | 616 (41%)           |
| Counties Manukau                      | 76 (41%)         | 48 (51%)          | 66 (35%)         | 214 (21.4%)         | 411 (28%)           |
| Waikato                               | 73 (39%)         | 11 (12%)          | 25 (13%)         | 349 (35.0%)         | 465 (31%)           |
| <b>Urban/rural</b>                    |                  |                   |                  |                     |                     |
| Urban                                 | 170 (92%)        | 92 (98%)          | 188 (99%)        | 887 (88.9%)         | 1,361 (91%)         |
| Rural                                 | 15 (8.1%)        | 2 (2.1%)          | 1 (0.5%)         | 111 (11.1%)         | 131 (8.8%)          |

\*Mothers' external prioritised ethnicity at baseline

Note: Maternal characteristics shown in Table 2 are at baseline



## 5. Summary

In addition to this technical report, users planning analyses of the GUINZ consented sub-set of data in the IDI should familiarise themselves with the original cohort profile and the cohort-alignment paper, relevant DCW reports, questionnaires, data dictionaries, and the Data User Guide (see Useful Resources). Together, these resources provide essential information on the original cohort design, available variables, and data collection methods, and will allow the user to gain an understanding of the differences between the full GUINZ dataset and the sub-set dataset available through the SNZ IDI.

For IDI users, a key consideration is that findings from the linked data should not be interpreted as automatically generalised to the same target population as analyses that could use the full GUINZ cohort. Assessment of topics like selection at each collection wave, attrition over time, analytic approaches, weighting, stratification, or sensitivity analyses are required for consideration of the research question.

## 6. Useful References

1. Morton, S. M. B., Atatoa Carr, P. E., Grant, C. C., Robinson, E. M., Bandara, D. K., Bird, A., Ivory, V. C., Kingi, T. K., Liang, R., Marks, E. J., Perese, L. M., Peterson, E. R., Pryor, J. E., Reese, E., Schmidt, J. M., Waldie, K. E., & Wall, C. (2013). Cohort profile: Growing Up in New Zealand. *International Journal of Epidemiology*, 42(1), 65–75. <https://doi.org/10.1093/ije/dyr206>
2. Morton, S. M. B., Ramke, J., Kinloch, J., Grant, C. C., Atatoa Carr, P., Leeson, H., Lee, A. C. L., & Robinson, E. (2015). Growing Up in New Zealand cohort alignment with all New Zealand births. *Australian and New Zealand Journal of Public Health*, 39(1), 82–87. <https://doi.org/10.1111/1753-6405.12220>
3. Growing Up in New Zealand. (2026) Data User Guide: March 2026. Auckland: *Growing Up in New Zealand*. The University of Auckland. [https://cdn.prod.website-files.com/684661c-0b478759e4c1bcb4a/69cb1b7f1f13d843fd68d264\\_DCW15\\_data\\_user\\_guide\\_March\\_2026.pdf](https://cdn.prod.website-files.com/684661c-0b478759e4c1bcb4a/69cb1b7f1f13d843fd68d264_DCW15_data_user_guide_March_2026.pdf)
4. Yao, E. S., Neumann, D., Taufa, S., Liang, R., Kingi, T. K., Langridge, F., & Paine, S. J. (2025). Ethnic-specific characteristics associated with longitudinal response patterns from prebirth to 12 years: evidence from Growing Up in New Zealand. *Journal of Epidemiology and Community Health*, 79(1), 19–26. <https://doi.org/10.1136/jech-2024-222428>
5. Morton, S. M. B., Atatoa Carr, P. E., Bandara, D. K., et al. (2010). Growing Up in New Zealand: A longitudinal study of New Zealand children and their families. Report 1: Before we are born. Auckland: *Growing Up in New Zealand*. <https://doi.org/10.17608/k6.auckland.25338901>
6. Pillai, A., Park, S. A., Napier, C., Edmonds, K., Prasad, R., & Paine, S. J. (2026). Now We Are 15 – Methods. Auckland: *Growing Up in New Zealand*, University of Auckland. <https://doi.org/10.17608/k6.auckland.31119892>
7. Napier, C., Yao, E., Prasad, R., Kedia, A., Fenton, D., Black, S., Pillai, A., Morton, S. M. B., & Paine, S. J. (2023). Now We Are 12: Introduction to the Growing Up in New Zealand 12-Year Data Collection Wave. Auckland: *Growing Up in New Zealand*. <https://doi.org/10.17608/k6.auckland.25999804>
8. Growing Up in New Zealand (2026). Growing Up in New Zealand: Now We Are 15. The University of Auckland. Collection. <https://doi.org/10.17608/k6.auckland.c.8312851>

### Suggested citation:

Growing Up in New Zealand (2026). Growing Up in New Zealand Data in the Statistics New Zealand Integrated Data Infrastructure – A technical document for data users v1.2 2026. Auckland: *Growing Up in New Zealand*.

© Growing Up in New Zealand 2026



Waipapa  
Taumata Rau  
University  
of Auckland

Growing Up  
in New Zealand