Linking 'data silos' to investigate anaemia among Aboriginal and Torres Strait Islander mothers and children in Far North Queensland

Dympna Leonard,[1] Petra Buettner,[1] Fintan Thompson,[1] Maria Makrides,[2] Robyn McDermott[1]

Abstract

Objective: Data collection 'silos' can be linked for health research. Anaemia in early childhood is a long-recognised health issue in remote Aboriginal communities of the Northern Territory and Western Australia, but information is lacking for Queensland. The objective of this work was to compile existing information from health and education data collections to investigate anaemia among Aboriginal and Torres Strait Islander mothers and their children in Far North Queensland.

Methods: Data mapping identified four health data collections and one education data collection holding relevant information. Data Custodians' approval was secured for release of linked de-identified information.

Results: Approval processes and preparation of the dataset for release took 23 months. Birth information was obtained for 2,205 mother–child pairs where the Aboriginal and/or Torres Strait Islander child was born in Far North Queensland between 2006 and 2010. Pathology information from before/during pregnancy was obtained for 2,126 mothers (96.4%), growth and haemoglobin information for 982 children (44.5%), and childhood development indicators at school entry for 963 children (43.7%).

Conclusion: Linking existing information 'silos' enables research into key public health issues.

Implications for public health: Information linkage is particularly valuable in respect of vulnerable populations including rural and remote Aboriginal and Torres Strait Islander peoples.

Key words: linkage, anaemia, Indigenous, mothers, children

Centralised data collections are potentially rich resources for public health research; however, their value is limited if the data is held in isolation from other relevant information sources.[1] The term 'data silos' has been used to describe such isolated data collections.[2] Linkage of data silos for longitudinal and intergenerational research can have particular advantages in respect of time and cost.[1,2] Here, we describe the process and results of our work to access and link existing data collections to investigate anaemia among Aboriginal and Torres Strait Islander mothers and their children in Far North Queensland.

Anaemia is a long-recognised problem among Aboriginal and Torres Strait Islander pre-school and school-aged children of remote communities in the Northern Territory and Western Australia.[3-7] Recent reports indicate that anaemia in pregnancy is also prevalent.[8] In remote Far North Queensland, most of the population (n=14,107 [71.5%]) is made up of Aboriginal and Torres Strait Islander people.[9] Similar issues with anaemia might be expected; however, there is currently no information to clarify the situation.

Anaemia is defined as low blood haemoglobin levels, measured in grams per litre (g/L). Cut-offs vary by age, sex and life stage and may be further adjusted for smoking and for locations at high altitude. The World Health Organization recommended cut-offs are the most commonly used (six months up to five years <110 g/L; 5–11years <115 g/L; 12–14 years <120 g/L; non-pregnant women 15 years and older <120g/L; pregnant women <110g/L, men 15 years and older <130 g/L).[10]

In many countries, including affluent countries such as Canada, the United States and Australia, anaemia is higher among the Indigenous populations compared with the general population.[11] This was shown in the recent Australian national health survey that identified participants who were at risk of anaemia.[12] Among Australian Aboriginal and Torres Strait Islander adults, 7.6% were at risk of anaemia, which is almost double the prevalence (4.5%) among non-Indigenous Australians (age adjusted rate ratio 1.9). More Australian Aboriginal and Torres Strait Islander women compared to men (10.3% vs. 4.8%) were at risk of anaemia and more Aboriginal and Torres Strait Islander people in remote compared to non-remote locations (10.1% vs. 6.9%).[12,13]

1. Centre for Chronic Disease Prevention, Australian Institute of Tropical Health and Medicine, College of Public Health, Medical and Veterinary Sciences, James Cook University, Queensland
2. Healthy Mothers, Babies and Children, South Australian Health Medical Research
Correspondence to: Ms Dympna Leonard, Centre for Chronic Disease Prevention, Australian Institute of Tropical Health and Medicine, James Cook University, PO Box 6811, Cairns, Queensland 4870; e-mail: dympna.leonard@jcu.edu.au
Submitted: January 2018; Revision requested: April 2018; Accepted: June 2018
The authors have stated they have no conflict of interest.

Aust NZ J Public Health. 2018; 42:456-62; doi: 10.1111/1753-6405.12821

The higher prevalence of anaemia among Aboriginal and Torres Strait Islander Australians is consistent with the lower socioeconomic status of Aboriginal and Torres Strait Islander people compared with other Australians.[14] Among all Australians, the risk of anaemia increases with lower incomes, and is higher among women than among men.[15]

Infection, inflammation and genetic conditions can cause anaemia.[16] However, deficiencies of iron and/or other nutrients remain the most common cause of anaemia among women and young children worldwide.[17] In the Northern Territory, the positive response to treatment with iron supplements for anaemic pre-school-aged and school-aged children indicates that iron deficiency is the cause of childhood anaemia in that setting.[7,18]

In early life, the principal source of iron for the rapidly growing infant is not breastmilk or infant formula but the iron endowed to the child by the mother, mostly during the last ten weeks of pregnancy.[19] Anaemia of a mother in pregnancy is strongly associated with early onset anaemia of her child.[20] For reasons that are not yet clear, maternal diabetes in pregnancy is also associated with early onset anaemia in the child.[21]

Anaemia has negative effects on the health of pregnant mothers, ranging from increased fatigue to increased risk of post-partum haemorrhage.[22] These effects, however, vary depending on the stage of gestation and severity of anaemia. Excessively high haemoglobin levels are also associated with poor pregnancy outcomes.[23] In young children, anaemia can compromise both health and development.[24,25] Anaemia in early childhood may have long-term negative effects, with lower levels of educational attainment during school years.[26,27] These detrimental effects can persist even when anaemia has been treated. Consequently, effective prevention of anaemia is important, especially in the first 1,000 days of life – through pregnancy to around two years of age – when growth and development are most rapid.[24,28]

The work described here has created linked records for mother–baby pairs from before pregnancy, through pregnancy, from birth, and through early childhood up to school entry using existing data collections. These linked intergenerational longitudinal records will be used to investigate anaemia among Aboriginal and Torres Strait Islander children and their mothers in Far North Queensland. This report describes the methods used to secure this information and the resultant data collection.

Methods

An overview of the planned research and the associated key variables by life-stage is shown in Table 1. Four Queensland Health data collections and the Department of Education Australian Early Development Census were identified as information sources. The key variables were identified from each respective data dictionary and listed in ethics and public health act applications.

Details of these five data collections are shown in Supplementary Table 1. Four are centralised whole-of-population data collections while one – Ferret (Ferret, Pen Computer Systems) – is used mainly in remote Far North Queensland.[29,30] The 38 localities using Ferret and the year when the Ferret system was rolled out are shown in Figure 1.

Scope

The geographic reach is Far North Queensland (Figure 1). The information was collected over 15 years, from 2000 to the end of 2015.

Study design and participants

The planned research will be a retrospective cohort study. Information was sought for two cohorts of Aboriginal and Torres Strait Islander mothers and their children.

The Cape York Child Growth cohort included Aboriginal and/or Torres Strait Islander children of the remote communities of Cape York, born between January 2006 and December 2008, and their mothers. These children were a subset of children included in previous unpublished health service research, born after the introduction of Ferret in the Cape York remote communities.

The 2009 and 2010 birth cohort included children and their mothers, where the child was born to an Aboriginal and/or Torres Strait Islander mother in Far North Queensland in 2009 or 2010.

Ethics and related approvals

Ethics approval granted by the Queensland Health Cairns and Hinterland Human Research Ethics Committee included a waiver of the requirement for participant consent to use their information. Details and timeframe for the subsequent Queensland Public Health Act approval processes are shown in Supplementary Table 2.

Data linkage

Once the required approvals had been secured, the requested information was extracted and provided to the Queensland Health data linkage team by the respective Data Custodians. The data linkage team created linkage keys for each mother and baby/child. Two of the four health service data collections (the Perinatal Data Collection and the Queensland Hospitals Admitted Patients Data Collection) are linked on an ongoing basis in a Master Linkage File that was accessed for this research.[30] The data linkage team used Linkage*Wiz* (Linkage*Wiz* v5.5.42 2015 Linkage*Wiz*Software http://www.linkagewiz.net/index.htm) for probabilistic linking of information from the three other data collections. Manual clerical review was also undertaken where required.[30]

Table 1: Sequence of life stage, key variables and data collection sources for each mother and baby pair required for planned research to investigate anaemia among Aboriginal and Torres Strait Islander mothers and their babies in Far North Queensland.

Life stages:	Mother (this pregnancy and prior) →	Baby at birth →	Child: birth to age 5 years →	Child: first year of school
Key variables	Ethnicity, location usual residence, parity/age at birth of cohort baby, anaemia in pregnancy, gestational diabetes, pregnancy inducted hypertension, pathology test results, height and weight, smoking, diet	Sex, gestational age, weight, length and head circumference at birth, APGAR 1&5, hospital admissions, length of stay, discharge status and ICD codes, initial infant feeding	Sequential measurements of weight, length/height and haemoglobin, early childhood development milestones (Y/N), hospitals admissions, length of stay, discharge status and ICD codes	Developmental Index for each of 5 domains: physical health and well being, social competence, emotional maturity, language and cognitive skills, communication skills and general knowledge. Developmental Categories: on track, at risk, vulnerable
Data sources	Auslab, Perinatal Data Collection (PDC)	Perinatal Data Collection (PDC), Queensland Hospitals Admitted Patients Data Collection (QHAPDC)	Ferret, Queensland Hospitals Admitted Patients Data Collection (QHAPDC)	Australian Early Development Census (AEDC)

Linkage files were given sequentially to the researchers as data became available and the final complete linkage file was provided in May 2017. The researchers used the given linkage keys to merge the information from the five data collections for each mother–child pair.

Assessment of data quality and of selection bias

For this report, data quality was assessed by considering the proportion of missing data and implausible values. Comparisons with census information on population numbers and ethnicity were made to assess data completeness and possible selection bias.

An additional comparison was made in respect of mothers and newborns of the 2009 and 2010 birth cohort. These mothers and babies came from both remote and non-remote localities in Far North Queensland. All of these babies had a record on the Perinatal Data Collection (PDC) but, as the Ferret system was used mainly in remote localities, a subset (n=728, 37.1%) of these children had a Ferret record as well as a PDC record. The mothers and babies where the child had a Ferret record were compared with the mothers and babies where the child did not have a Ferret record. These comparisons were made to assess if relying on information from the Ferret data system would introduce any systematic bias in the subsequent analysis.

Definitions of key characteristics

Key characteristics of the mothers, babies and young children are presented here to describe the information obtained; for example, body mass index of mothers, prematurity of babies. For mothers, conditions in pregnancy (anaemia, gestational diabetes and pregnancy-induced hypertension) are as reported in the Perinatal Data Collection (PDC). Other definitions are those used by the National Health and Medical Research Council, the Australian Institute of Health and Welfare and the Australian Bureau of Statistics. For details of the definitions used, see Supplementary Table 3. World Health Organization (WHO) guidelines for use of the WHO Child Growth Standards were followed.[31,32] Weight and height measurements that resulted in weight for age or height for age z-scores of less than -6 or greater than +5 (weight for age) or +6 (length/height for age) were considered implausible values.[31,33]

Results

The process of approvals, data extraction and preparation of the linked data took 23 months from ethics approval. Six interim releases were made as data became available, at the researchers' request. Supplementary Table 2 provides more detail. The information secured for each of the two cohorts of mothers and their children is summarised below.

The Cape York Child Growth Cohort 2006 to 2008

Perinatal data records for birth were provided for 380 children born in 2006, 2007 or 2008 and their mothers (n=339); 87.6% of the 434 children included in previous research. To ensure independence of events for subsequent statistical analysis, children who were not the first birth to a mother during those three years (n=40) and second-born twins (n=3) were excluded, leaving records for 337 unique mother–child pairs. The process of exclusions and results of linkage with other data collections are shown in Figure 1.

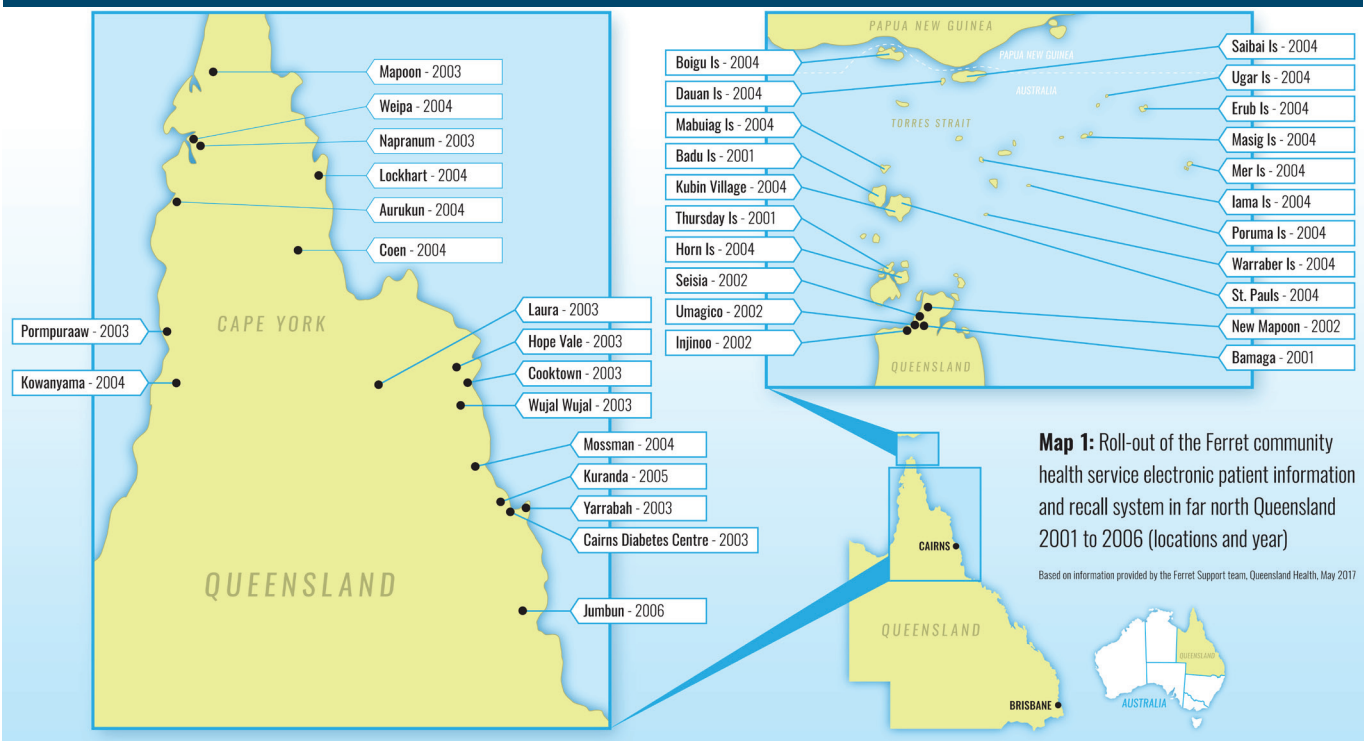
Key descriptive information for these 337 mothers and their babies at birth is shown in Table 2.

The 2009 and 2010 birth cohort

Perinatal Data Collection records for birth were provided for 2,167 babies born to 1,993 Aboriginal and/or Torres Strait Islander mothers in 2009 or 2010 in Far North Queensland. Of these, 20 babies were stillborn and 11 died as neonates, leaving 2,136 surviving babies.

Children who were the second baby (n=154) or third baby (n=1) born to the same mother in that two-year period and 16 babies who

Figure 1: The rollout of Ferret in Far North Queensland 2001 to 2006.



were the second-born of twins are excluded from this report, leaving information for 1,965 unique mother–child pairs. The process of exclusions and results of linkage with the other data collections are shown in Figure 2. Descriptive information for these mothers and their babies (n=1,965) and for the subset of mothers and babies (n=728) with longitudinal information from the Ferret system is shown in Table 2.

Data quality and completeness

The quality and completeness of the data provided varied between different data sources with different variables, as shown in Table 3. Information from the Perinatal Data Collection (PDC) was complete for some variables (mothers' dates of birth, the date of birth and sex of babies, birth weight, plurality and method of birth). More information was missing in respect of mothers' weights, heights and parity. The proportion of missing PDC information reduced over time. Pathology measurements of haemoglobin were available for most mothers (Cape York mothers, 87.8%; 2009 and 2010 birth cohort mothers 97.0%) and, to a lesser extent, measurements of glucose tolerance and iron status. For most mothers (between 73.5% and 85.9%) information on folate and vitamin B_{12} levels was not available.

Comparisons with census information

Data completeness was also assessed by comparison of child numbers and information on the ethnicity of mothers with census information for remote areas.

Cape York Child Growth cohort

For the Cape York communities, Census 2006 results show 427 resident children who were born in the three years preceding the August 2006 Census. This number is close to the 434 children included in the previous research project, although somewhat higher than the 380 children for whom a record of birth between 2006 and 2008 was located on the Perinatal Data Collection.

2009 and 2010 birth cohort

The locations where the Ferret system was used, as shown in Figure 1, were mainly in Cape York (n=12) and in the Torres Strait (n=21). When Census 2011 population figures are combined for Torres Strait and Cape York, the total population of Aboriginal and Torres Strait Islander children under five years of age is 1,830. On a pro-rata basis, this is equivalent to 732 children who were born in any two-year period from Census 2006 to Census 2011.[34] This figure is close to the number of children (n=728) identified by this research who were born in 2009 and 2010 and who had both Perinatal Data Collection and Ferret records.

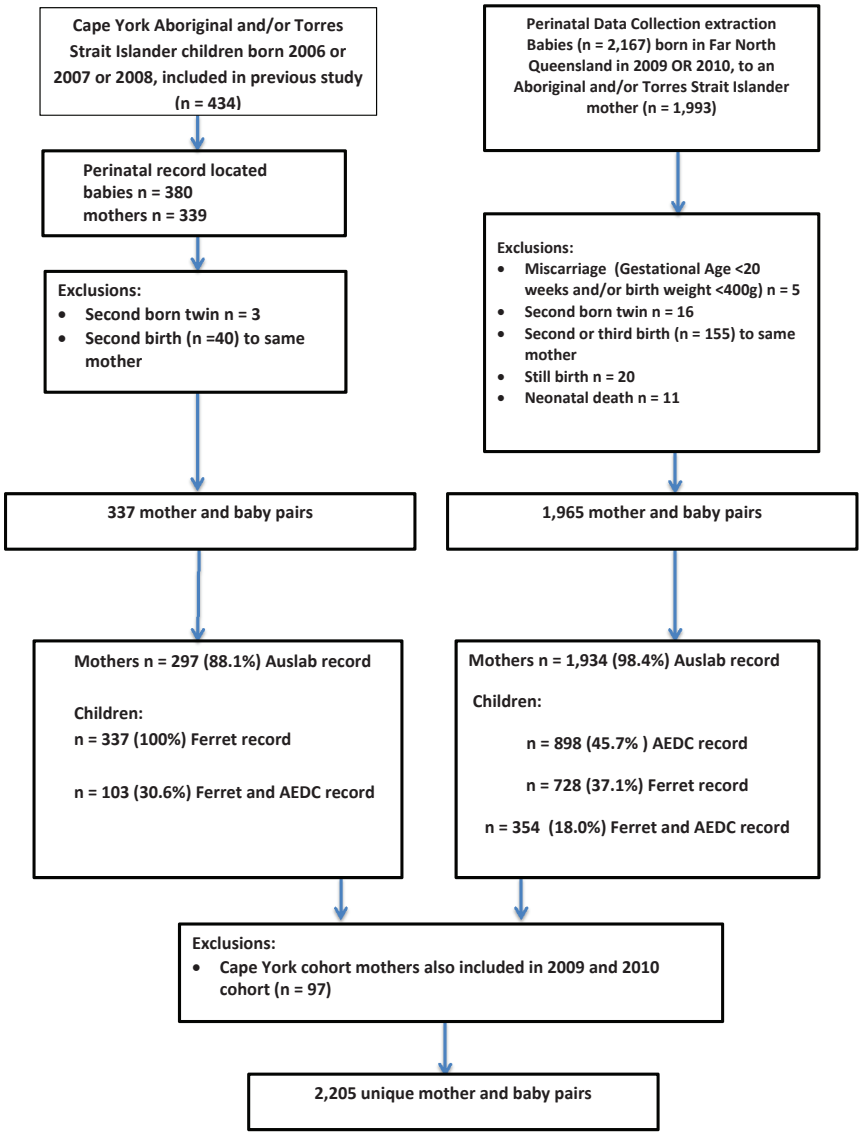
Information in respect of ethnicity of mothers is shown in Supplementary Figures 1a and 1b. The ethnicity of the mothers from Cape York and the Torres Strait who were included in both cohorts is consistent with information on ethnicity of residents of Cape York and the Torres Strait as reported in the 2011 Census.

Comparing mothers and babies with a child Ferret record and those without a child Ferret record

Differences for mothers were found in respect of residence in a remote location (PDC only: 7.8% remote residence vs. PDC and Ferret: 86.3% remote residence; $p<0.001$), in ethnicity and in gestational diabetes in pregnancy. Mothers whose child had a Ferret record were more likely to be Torres Strait Islander (51.1% vs. 24.8%) and less likely to be Aboriginal (37.4% vs. 61.9%; $p<0.001$). More mothers whose child had a Ferret record had gestational diabetes (12.1% vs. 9.2%, $p=0.043$). No other significant difference was seen in respect of the mothers or their babies at birth. The results of this comparison are shown in Supplementary Table 4.

Subsequent merging of information for the two cohorts of mothers and babies identified 97 mothers in the 2009–2010 birth cohort (n=1,965), who were already included in the earlier Cape York Child Growth cohort. Excluding these 97 duplicate mothers and their later babies reduced the numbers

Figure 2: Flow diagram – data provided and exclusions for both cohorts.



of mother–baby pairs in the 2009–2010 cohort to 1,868 pairs. In total, therefore, this work assembled intergenerational health information for 2,205 Aboriginal and Torres Strait Islander mother–baby pairs from Far North Queensland where the child was the first child born to that mother between 2006 and 2010.

Discussion

The work described here has resulted in a dataset with longitudinal intergenerational information for 2,205 Aboriginal and Torres Strait Islander mother–child pairs from prior to pregnancy, through pregnancy, from birth and through early childhood to the first year of school, recorded over a period of 15 years. The process of obtaining these data was lengthy (23 months) but was much less time than the 15-plus years required for a prospective study with an equivalent timespan.

The release of earlier versions, as each stage of linkage was completed, enabled the researchers to become familiar with the dataset. This process allowed gaps and errors to be identified and rectified. Preparatory analysis – of child growth parameters, for example – was undertaken prior to the release of the completed linkage file.

Comparisons with census information showed consistency in respect of child participant numbers and the ethnic mix of mothers from remote localities.[34] Differences in regional boundaries used by various government entities meant it was difficult to make similar comparisons for non-remote locations. However, the comparison between those mothers and babies where the child had a subsequent Ferret record and those where the child did not have a Ferret record was effectively a comparison of remote and non-remote participants. These comparisons found few differences, apart from remote residence, ethnicity and diabetes in pregnancy. These findings reflect the high Ferret coverage in the Torres Strait where the incidence of diabetes in pregnancy is particularly high.[35]

The researchers will analyse the data to investigate anaemia among these Aboriginal and Torres Strait Islander mothers and their children. Risk factors for early childhood anaemia that relate to the health of mothers will be explored (age, parity, anaemia, iron status and glucose tolerance) and factors relating to the child (birth weight, gestational age, early infant feeding, rate of growth). Information on early childhood development indicators at school entry will be used to assess the consequences of early childhood anaemia.

Table 2: Key Characteristics of each cohort - Mothers and their babies at birth, including the subset of the 2009 and 2010 birth cohort where the child had a Ferret longitudinal record of growth and haemoglobin measurements.

	Cape York Child Growth	All 2009 & 2010 Births	2009 & 2010 Births with Ferret record
	Mothers (n=337)	**Mothers (n=1,965)**	**Mothers (n=728)**
Ethnicity			
Aboriginal	286 (84.9%)	1,038 (52.8%)	272 (37.4%)
Torres Strait Islander	18 (5.3%)	679 (34.6%)	372 (51.1%)
Aboriginal and Torres Strait Islander	18 (5.3%)	248 (12.6%)	84 (11.5%)
Non-Indigenous	15 (4.5%)	–	–
Location usual residence			
Torres & NPA	3 (0.9%)	443 (22.5%)	375 (51.5%)
Cape York	298 (88.4%)	304 (15.5%)	270 (37.1%)
Other FNQ	26 (7.7%)	1,199 (61.0%)	78 (10.7%)
not FNQ	10 (3.0%)	19 (1.0%)	5 (0.7%)
Age years mean (SD) range	24.9 (6.4) 15–41	25.3 (6.4) 13–48	25.0 (6.2) 13–48
Parity median (range)	2 (0–8)	2 (0–16)	2 (0–10)
Body Mass Index (BMI) (kg/m2)	n=122	n=1,834	n=679
mean (SD), range	23.7 (5.8), 14.9–37.7	27.1 (6.6), 14.3–55.9	27.4 (6.8), 14.3–55.9
Body Mass Index categories			
Under-weight	28 (23.0%)	114 (6.2%)	51 (7.5%)
Healthy weight	46 (37.7%)	684 (37.3%)	225 (33.1%)
Over-weight	31 (25.4%)	456 (24.9%)	167 (24.6%)
Obese	17 (13.9%)	580 (31.6%)	236 (34.8%)
Perinatal Data Collection (PDC) - pregnancy information			
Anaemia	14 (4.2%)	75 (3.8%)	28 (3.9%)
Gestational Diabetes	11 (3.6%)	202 (10.3%)	88 (12.1%)
Pregnancy Induced Hypertension	27 (7.4%)	96 (4.9%)	35 (4.8%)
Smoking	215 (64.6%) (n=333)	1,113 (56.8%) (n=1,960)	426 (58.9%) (n=726)
	Babies (n=337)	**Babies (n=1,965)**	**Babies (n=728)**
Boys/Girls	51.3%/48.7%	54.0%/46.0%	52.9%/47.1%
Gestational Age weeks median, range	39, 27–42	39, 22–43	39, 26–42
Premature n (%) (95% CI)	40 (11.9%) (8.4%, 15.3%)	209 (10.7%) (9.3%, 12.0%)	76 (10.4%) (8.2%, 12.7%)
Birth Weight[a] grams mean (SD) range	3,089 (602.3) 800–5,320	3,247.1 (629) 495*–5,430	3,273 (600) 960–5,050
Low birth weight (<2,500g)	45 (13.4%)	196 (10.0%)	65 (8.9%)
High birth weight (>= 4,000g)	15 (4.5%)	185 (9.4%)	62 (8.5%)

Notes:

a: baby with birth weight 495g was recorded as a live birth - there is nil record of neonatal death for this baby

There were some issues of data quality and completeness, although reductions in missing values over time indicate ongoing quality improvement. For the analysis, we will use the STATA statistical package (Stata version 13, StataCorp, Lakeway Drive, College Station, Texas) to enable us to conduct multivariable analysis, with and without data imputation, to gauge the impact of missing information on the outcome measures.[36]

An inquiry by the Productivity Commission into the use of existing data collections recommended greater transparency and changes in the legal framework to increase accessibility, which may shorten the time required for approval processes in future.[37] However, access to existing data collections entails ethical and legal considerations including issues of privacy and confidentiality that require time for proper consideration. This is particularly true for research relating to Aboriginal and Torres Strait Islander peoples.[38]

The recently endorsed Australian National Digital Health Strategy should provide a framework for integrated health service data systems, replacing the current data silos.[2] In Far North Queensland, there has been an increase in community-controlled health service providers in recent years. An integrated health service data system will have benefits for service provision with

this welcome increased diversity of service providers – and benefits for future data linkage for research.

Implications for public health

The work described in this report has secured a dataset of linked information from four health data collections and one education data collection, which will be a valuable resource in investigating the issue of anaemia among Aboriginal and Torres Strait Islander mothers and their children in Far North Queensland. This report illustrates how the linkage of existing data resources can provide intergenerational information for health research. However, the true value of the resultant data collection will be demonstrated by the subsequent planned research and reporting. To use a nutrition-related analogy, the proof of the pudding will be in the eating.

Acknowledgements

We would like to acknowledge and thank the Aboriginal and Torres Strait Islander leaders of the key community-controlled health service organisations in Far North Queensland who considered and endorsed the proposed research, providing the support that made this research possible.

In addition, we acknowledge and thank the Data Custodians and their research and data management staff.

We especially want to acknowledge the contributions of the following people:

- Staff of the Queensland Health Statistical Services Branch, with particular thanks to Rachel Garry for support for Public Health Act applications and variations, and to Dale Steinhardt for data extraction and linkage.
- Members of the Queensland Health Ferret Support Team, with particular thanks to David Woodman for data extraction, to Vikki Tierney, Paula Lush and Cherie Blofield for information on the rollout of Ferret, and to Rosemary Schmidt for support for Public Health Act applications.
- Staff of the Australian Early Development Census support team, with particular thanks to Irene Adam-Manik for support for the AEDC application and data extraction.
- Staff of the Queensland Pathology Clinical Information Systems Support Unit, with particular thanks to Wanda Sprenger for data extraction.

Funding

Dympna Leonard was supported by the Australian National Health and Medical Research Council postgraduate scholarship APP1092732. Apart from this funding support, the National Health and Medical Research Council did not have any role in respect of this manuscript.

Table 3: Data quality – missing, implausible, and available data – examples from four[a] data collections.

	Cape York Child Growth Born 2006 – 2008 n=337 mother and baby pairs n=337 with Ferret longitudinal record n (%)	2009 & 2010 Births n=1,965 mother and baby pairs n=728 with Ferret longitudinal record n (%)
Perinatal Data Collection – missing information		
Mothers		
Ethnicity	nil	nil
Location usual residence	nil	nil
Date of birth	nil	nil
Smoking in pregnancy	4(1.2%)	5 (0.3%)
Parity	109 (32.3%)	526 (26.8%)
Weight	190 (56.4%)	79 (4.0%)
Height	196 (58.2%)	117 (6.0%)
Weight and/or Height	204 (60.5%)	136 (6.9%)
Babies – missing information		
Birth weight	nil	1 (0.05%)
Gestational age at birth	nil	5 (0.3%)
Sex	nil	nil
Ferret – child records		
Inconsistent ethnicity records	nil	9 (1.2%)
Date of measurement prior to date of birth	nil	15 (0.2%, n=8,539)
Implausible weight measurements	170 (2%, n=8,328 measurements)	45 (0.6%, n=7,150 measurements)
Implausible lengths/heights	28 (2.3%, n=1,202 measurements)	45 (1.7%, n=2,622 measurements)
Auslab data missing for mothers		
Haemoglobin record in cohort pregnancy	41 (12.2%)	59 (3.0%)
Iron status record for mother in cohort pregnancy	125 (37.1%)	752 (38.3%)
Glucose tolerance before or during pregnancy	31 (10.3%)	583 (29.7%)
Folate measurement before or during pregnancy	277 (73.5%)	1,552 (79.0%)
Vitamin B12 measurement before or during pregnancy	324 (85.9%)	1,603 (81.6%)
Australian Early Development Census record available (child) – 2012	103 (30.6%)	n/a
Australian Early Development Census record available (child) – 2015	n/a	898 (45.7%)

Notes:

a: Only some children would have been admitted to hospital, so it is not possible to ascertain if Queensland Hospital admission records are missing

References

1. Buttner P, Muller R. *Epidemiology*. 2nd ed. Melbourne (AUST): Oxford University Press; 2015.
2. Australian Digital Health Agency. *Australia's National Digital Health Strategy*. Canberra (AUST): ADHE; 2017.
3. Gracey MS, Sullivan H. Growth of Aboriginal infants in the first year of life in remote communities in north-west Australia. *Ann Hum Biol*. 1988;15(5):375-82.
4. Brewster DR. Iron deficiency in minority groups in Australia. *J Paediatr Child Health*. 2004;40:422-3.
5. Smith RM, Smith PA, McKinnon M, Gracey M. Birthweight and growth of infants in five Aboriginal communities. *Aust N Z J Public Health*. 2000(24):124-35.
6. Bar-Zeev SJ, Kruske SG, Barclay LM, Bar-Zeev N, Kildea SV. Adherence to management guidelines for growth faltering and anaemia in remote dwelling Australian Aboriginal infants and barriers to health service delivery. *BMC Health Serv Res*. 2013;13:250.
7. Udovicich C, Perera K, Leahy C. Anaemia in school-aged children in an Australian Indigenous community. *Aust J Rural Health*. 2017;25(5):285-9.
8. Bar-Zeev S, Barclay L, Kruske S, Kildea S. Factors affecting the quality of antenatal care provided to remote dwelling Aboriginal women in northern Australia. *Midwifery*. 2014;30(3):289-96.
9. Queensland Government Statistician's Office. *Resident Population Profile*. Brisbane (AUST): QGSO; 2017.
10. World Health Organisation. *Haemoglobin Concentrations for the Diagnosis of Anaemia and Assessment of Severity*. Geneva (CHE): WHO; 2011.
11. Khambalia AZ, Aimone AM, Zlotkin SH. Burden of anemia among Indigenous populations. *Nutr Rev*. 2011;69(12):693-719.
12. Australian Bureau of Statistics. *Australian Aboriginal and Torres Strait Islander Health Survey: Biomedical Results 2012-13*. Canberra (AUST): ABS; 2014.
13. Australian Bureau of Statistics. *Australian Health Survey: Biomedical Results for Chronic Disease, 2011-12 – Anaemia*. Canberra (AUST): ABS; 2013.
14. Australian Institute of Health and Wellfare. *The Health and Welfare of Australia's Aboriginal and Torres Strait Islander Peoples*. Canberra (AUST): AIHW; 2015.
15. Callander EJ, Schofield DJ. Is there a mismatch between who gets iron supplementation and who needs it? A cross-sectional study of iron supplements, iron deficiency anaemia and socio-economic status in Australia. *Br J Nutr*. 2016;115(4):703-8.

16. Strengthening Partnerships, Results, and Innovations in Nutrition Globally. *Changing the Way We Think about Micronutrient Assessment and Anemia Programming. Findings from the Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) Project*. Arlington (VA): SPRING; 2017.
17. Stoltzfus RJ, Klemm R. Research, policy, and programmatic considerations from the Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr*. 2017;106 Suppl 1:428-34.
18. Kruske SG, Ruben AR, Brewster DR. An iron treatment trial in an Aboriginal community: Imporving non-adherence. *J Paediatr Child Health*. 1999;35(2):153-8.
19. Dewey KG, Chaparro CM. Iron status of breast-fed infants. *Proc Nutr Soc*. 2007;66(3):412-22.
20. Balarajan Y, Ramakrishnan U, Özaltin E, Shankar AH, Subramanian SV. Anaemia in low-income and middle-income countries. *Lancet*. 2011;378(9809):2123-35.
21. Rao R, Georgieff MK. Iron in fetal and neonatal nutrition. *Semin Fetal Neonatal Med*. 2007;12(1):54-63.
22. Allen LH. Anaemia and iron deficiency: Effect on pregnancy outcome. *Am J Clin Nutr*. 2000;71 Suppl:1280-7.
23. Dewey KG, Oaks B. U-shaped curve for risk associated with maternal hemoglobin, iron status, or iron supplementation. *Am J Clin Nutr*. 2017;106 Suppl 6:1694-1702.
24. Prado EL, Dewey KG. Nutrition and brain development in early life. *Nutr Rev*. 2014;72(4):267-84.
25. Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet*. 2013;382(9890):427-51.
26. Grantham-McGregor S, Baker-Henningham H. Iron Deficiency in childhood: Causes and consequences for child development. *Annales Nestlé Eng*. 2010;68(3):105-19.
27. Lozoff B. Iron deficiency and child development. *Food Nutr Bull*. 2007;28(4):S560-71.
28. Georgieff MK. Iron assessment to protect the developing brain. *Am J Clin Nutr*. 2017;106 Suppl 6:1588-93.
29. Queensland Health. *Expanded Model of Primary Health Care*. Cairns (AUST): State Government of Queensland; 2005.
30. Queensland Health. *Queensland Data Linkage Framework*. Brisbane (AUST): State Government of Queensland; 2015.
31. World Health Organization. *WHO Child Growth Standards: I Growup Package Manual*. Genieva (CHE): WHO; 2006.
32. World Health Organization. *WHO Child Growth Standards*. Geneva (CHE): World Health Organization; 2006.
33. World Health Organization Multicentre Growth Reference Study Group. Chpt 7: Computation of Centiles and Z-Scores. In: *WHO Child Growth Standards: Methods and Development*. Geneva (CHE): WHO; 2006.
34. Australian Bureau of Statistics. *Census*. Canberra (AUST): ABS; 2011.
35. Falhammar H, Davis B, Bond D, Sinha AK. Maternal and neonatal outcomes in the Torres Strait Islands with a sixfold increase in type 2 diabetes in pregnancy over six years. *Aust N Z J Obstet Gynaecol*. 2010;50(2):120-6.
36. Enders CK. *Applied Missing Data Analysis*. New York (NY): Guilford Press; 2010.
37. Productivity Commission. *Data Availaibility and Use*. Canberra (AUST): Government of Australia; 2017.
38. National Health and Medical Research Council. *Values and Ethics: Guidelines for Ethical Conduct in Aboriginal and Torres Strait Islander Health Research*. Canberra (AUST): NHMRC; 2003.

Supporting Information

Additional supporting information may be found in the online version of this article:

Supplementary Table 1: Data Collections used to source information for planned research to investigate maternal and early childhood anaemia.

Supplementary Table 2: Time frame for Queensland Public Health Act approvals, data linkage and release.

Supplementary Table 3: Definitions of variables .

Supplementary Table 4: 2009 and 2010 birth cohort: comparing characteristics of mothers and babies where the child had a Perinatal Data Collection (PDC) record only compared to mothers and babies where child had both a PDC and a Ferret record.

Supplementary Figure 1a: Comparison with Census 2011 – ethnicity of mothers from Cape York.

Supplementary Figure 1b: Comparison with Census 2011 – ethnicity of mothers from Torres Strait.