publichealth research&practice

November 2014; Vol. 25(1):e2511407

doi: http://dx.doi.org/10.17061/phrp2511407

www.phrp.com.au

Research – Methods

Approaches to maximising the accuracy of anthropometric data on children: review and empirical evaluation using the Australian Longitudinal Study of Indigenous Children

Katherine A Thurber[a,b], Emily Banks[a] and Cathy Banwell[a]

[a] National Centre for Epidemiology and Population Health, Australian National University, Canberra, ACT, Australia

[b] Corresponding author: Katherine.thurber@anu.edu.au

Article history

Publication date: November 2014

Citation: Thurber K, Banks E, Banwell C. Approaches to maximising the accuracy of anthropometric data on children: review and empirical evaluation using the Australian Longitudinal Study of Indigenous Children. Public Health Res Practice. 2014;25(1):e2511407. doi: http://dx.doi.org/10.17061/phrp2511407

Key points

- No anthropometric data are 100% accurate, even when 'gold standard' methods are employed
- However, existing published methods for improving the accuracy of anthropometric data seem to be ad hoc or lacking in clear justification
- Based on current evidence, we developed a probabilistic approach to identifying anthropometric measurements that have an increased likelihood of being implausible
- This method was applied to data from the Australian Longitudinal Study of Indigenous Children, and is intended to be reproducible in other settings

Abstract

Aim: Despite the burgeoning research interest in weight status, in parallel with the increase in obesity worldwide, research describing methods to optimise the validity and accuracy of measured anthropometric data is lacking. Even when 'gold standard' methods are employed, no data are 100% accurate, yet the accuracy of anthropometric data is critical to produce robust and interpretable findings. To date, described methods for identifying data that are likely to be inaccurate seem to be ad hoc or lacking in clear justification.

Study type: Methods.

Methods: This paper reviews approaches to evaluating the accuracy of cross-sectional and longitudinal data on height and weight in children, focusing on recommendations from the World Health Organization (WHO). This review, together with expert consultation, informed the development of a method for processing and verifying longitudinal anthropometric measurements of children. This approach was then applied to data from the Australian Longitudinal Study of Indigenous Children.

Results: The review identified the need to assess the likely plausibility of data by (a) examining deviation from the WHO reference population by calculating age- and sex-adjusted height, weight and body mass index z-scores, and (b) examining changes in height and weight in individuals over time. The method developed identified extreme measurements and implausible intraindividual trajectories. It provides evidence-based criteria for the exclusion of data points that are most likely to be affected by measurement error.

Conclusions: This paper presents a probabilistic approach to identifying anthropometric measurements that are likely to be implausible. This systematic, practical method is intended to be reproducible in other settings, including for validating large databases.

Introduction

For research to produce meaningful findings, the underlying data need to be accurate[1]; however, no data are 100% accurate, even when 'gold standard' methods are used. The collection of data, particularly longitudinal data, is expensive. Therefore, it is prudent to maximise the benefit from using these data by maximising their accuracy – including using appropriate measurement techniques and carefully processing collected data.

Despite the burgeoning research interest in obesity[2], research describing methods to assess the accuracy of measured anthropometric data is lacking.[3,4] Inaccuracies in anthropometric data can result from equipment, measurement, recording and data entry error, and these errors can alter the interpretation of an individual's weight status.[4] Without rigorous methods to underpin the accuracy of anthropometric measurements, research that uses these data – such as evaluation of weight-loss interventions – is likely to be unduly affected by measurement error. However, described methods for identifying data that are likely to be inaccurate seem to be ad hoc, based predominantly on convention and/or lacking clear justification.

Most defined methods for cross-sectional data rely on the exclusion of weights or heights outside a pre-specified range[5,6], but approaches are inconsistent, and the reason for, and clinical significance of, the chosen cut-off points is not usually explained. Although this type of approach may be appropriate for adults, it is more difficult to employ for children, because the plausible range for height and weight varies widely with age.[3] Evaluation of longitudinal data involves additional complexity, particularly for children, because intraindividual variation in height and weight that can be attributed to normal growth needs to be differentiated from implausible variation that is more likely a result of measurement error. A few longitudinal studies mention the use of methods to evaluate the plausibility of intraindividual changes in height or weight status over time[1,7–9]; however, we have been unable to locate an explicit description of these processes.

Because it is not ultimately possible to determine the true value of a measurement, determining the accuracy of any measurement is problematic.[4] It may be possible to re-measure all children with questioned measurements in studies with small samples, but this is not practical for large samples. This approach would still require a method to identify the measurements requiring review. An alternative to this resource-intensive re-measuring method is a systematic approach to identifying measurements that are more likely to be a result of measurement error than a true representation of extreme height or weight. These a priori, agreed (not post-hoc, data-derived) methods could be applied to both measurements in cross-sectional data and intraindividual changes in longitudinal data.

This study aimed to (a) review the approaches for evaluating the accuracy of cross-sectional and longitudinal anthropometric data, (b) interview data collection officers about the difficulties with measuring height and weight, (c) devise an evidence-based approach for the probabilistic evaluation of measurement accuracy informed by (a), (b) and expert advice, and (d) apply and empirically evaluate this approach to cross-sectional and longitudinal anthropometric data in the Longitudinal Study of Indigenous Children (LSIC).

Methods

Literature review

Literature was reviewed to identify published methods for evaluating the accuracy of anthropometric data on children. The initial search focused on World Health Organization (WHO) documentation, as WHO is the leading expert group in the field. Although WHO has published methods for identifying 'implausible' anthropometric measurements in cross-sectional data, it does not explicitly describe processes to identify 'implausible' intraindividual changes.

The literature search was expanded to include longitudinal studies that collected anthropometric data, and because of the limited research in the area, we examined studies of both children and adults. This search focused particularly on methods used by the Longitudinal Study of Australian Children (LSAC), a study similar to LSIC in design and conceptual framework. The methods to evaluate anthropometric data used by LSAC could present a solid template for use with LSIC.

LSIC study design

LSIC is a cohort study of up to 1759 Aboriginal and Torres Strait Islander children living across Australia. The survey is managed by the Australian Government Department of Social Services (DSS) and funded by the Australian Government.[10] Indigenous Research Administration Officers (RAOs) conducted structured interviews annually with the study children, a primary carer and a secondary carer. Carers reported on their child's general health, and RAOs measured children's height and weight. Despite the demonstrated health inequity between Indigenous and non-Indigenous Australians, LSIC is the first national longitudinal study to examine the life-course development of Indigenous children. Thus, these data have the potential to fill a large research gap.

LSIC anthropometric data

RAOs sought permission from all interviewed parents and carers to measure their child's height and weight at each interview. To help ensure a correct recording, RAOs were trained to take each measurement three times. Homedics digital scales (model SC-305-AOU-4209),

which are accurate to 100 grams, were used to weigh children. In the first wave of interviews, RAOs used plastic height-measuring sticks to measure children's standing height and tape measures to measure small infants' recumbent length; this equipment was chosen for ease of transport to the most remote locations (via small aircraft and boats with weight restrictions). To improve the quality of data collected in later waves, RAOs used Soehnle stadiometers (professional model 5003), which are accurate to the nearest millimetre.

If carers were not comfortable having the RAOs take these measurements, they were invited to take the measurements themselves or to report the most recent measurement of their child's height and weight recorded in their child's health record book ('baby book'), which were taken by health professionals in a controlled setting.

Carers were increasingly willing to have children measured as the study progressed, and the proportion of height and weight measurements taken from the baby book (rather than measured directly) dropped from 3% and 6%, respectively, to less than 1% by the fourth interview (see Table 1).

The accuracy of the height and weight measurements collected in LSIC required evaluation before the data could be released for researchers' use. Although every effort was made to collect accurate data, the interview context was often inhibiting. RAOs and members of the LSIC team acknowledged the difficulty in taking these measurements, especially when children were unable to stand still while being measured or when flat surfaces were not available for measuring. The team recognised that data integrity might be compromised, so a process was developed to remove data that were likely to be inaccurate.

Table 1. Number of children with weight and height recorded in each wave

	Wave 1	Wave 2	Wave 3	Wave 4
Number of children interviewed	1671	1523	1404	1283
Height recorded (% of children interviewed)	1322 (79.11)	1415 (92.91)	1318 (93.87)	1245 (97.04)
From baby book (% of all recorded heights)	38 (2.87)	12 (0.85)	18 (1.37)	5 (0.40)
Measured by RAO (% of all recorded heights)	1284 (97.13)	1403 (99.15)	1300 (98.63)	1240 (99.60)
Weight recorded (% of children interviewed)	1365 (81.69)	1451 (95.27)	1334 (95.01)	1257 (97.97)
From baby book (% of all recorded weights)	77 (5.64)	59 (4.07)	17 (1.27)	8 (0.64)
Measured by RAO (% of all recorded weights)	1288 (94.36)	1392 (95.93)	1317 (98.73)	1249 (99.36)
Both height and weight recorded (% of children interviewed)	1304 (78.04)	1408 (92.45)	1308 (93.16)	1245 (97.04)

RAO = Research Administration Officer
Source: Thurber[11], The Longitudinal Study of Indigenous Children[12]

Factors impacting on height and weight measurement in LSIC

Before the height and weight data were evaluated, RAOs provided insight into the potential barriers hindering the collection of accurate data. Individual interviews and a focus group discussion were held with eight RAOs, representing 73% of those currently employed (see Thurber[11] for details). RAOs expressed concerns and difficulties relating to measuring children, such as technological limitations and the need to develop a relationship of trust with participants. Overall, the RAOs believed the accuracy of anthropometric data collection had improved over the course of the study. For example, RAOs stated that the use of the Soehlne stadiometer, starting in wave 2, increased their ability to take precise measurements.[11]

These interviews highlighted that the context in which these measurements were taken was particularly important. In some cases, measurements had to be taken outside on uneven ground, and in other cases, measurements were taken while the study child's siblings were present and distracting the study child. In some cases, RAOs recorded on the survey tools the conditions interfering with the measurement process, indicating the decreased reliability of those measurements. However, the recording of descriptive comments was not universal, so the circumstances in which measurements were taken, and the impact on data quality, was not always known. This made clear the need for a method to identify attributes of data that indicate that they were likely to be affected materially.

Standardisation of height and weight

Z-scores for height, weight and body mass index (BMI; calculated as weight divided by height squared) were calculated to determine the difference between a child's measurement and the median measurement of children of the same age and sex in the WHO Multicentre Growth Reference Study (a sample of 8440 healthy, breastfed infants from six countries).[13]

Cut-off points based on the statistical distribution of the reference are used to identify z-scores that are considered to be in the 'normal' range. Normal height and weight z-scores are classified as those falling in the middle 95% of the reference distribution, with a z-score between –2 and +2; z-scores in the lowest 2.5% or highest 2.5% are considered low or high, respectively.[14,15]

These values, though outside the healthy normal range, are still plausible.

BMI cut-offs are used to indicate underweight, normal weight, overweight and obesity. More conservative cut-off points are used for younger children because the health impact of excess weight at younger ages is less certain.[16] For children under five years of age, a BMI z-score of +1 indicates that a child is at risk of overweight, a z-score of +2 indicates that a child is overweight and a z-score of +3 indicates that a child is obese.[16] For children over five years of age, overweight is defined as a z-score exceeding +1, and obesity as a z-score exceeding +2. For both age groups, a z-score lower than –2 indicates underweight.[17]

Results

Methods for identifying implausible data described in the literature

Cross-sectional data

Most cross-sectional studies use cut-off points for implausible data based on raw height and weight values. For example, Das and colleagues define implausible measurements for adults as height values outside 122–213 cm and weight values outside 75–500 lb (34–226.5 kg).[5] Kahwati et al. use the same cut-offs for height values, and exclude weight values outside 70–700 lb (31.5–317.5 kg).[6]

In contrast, WHO provides guidelines for excluding data based on 'extreme' z-scores.[18] The use of z-score–based cut-off points allows the range of plausible height and weight values to vary with age, accommodating the wide variation observed in childhood growth patterns. Further, examination of BMI z-scores allows the plausibility of the combination of height and weight for a child at any given age to be considered.

The plausibility of a measurement decreases with the increasing magnitude of its z-score, and the probability that the measurement resulted from measurement or recording error increases. It is necessary to determine the point at which the probability that a measurement represents true deviation from the reference median is lower than the probability that the measurement is an error. WHO has defined a range of values that are biologically plausible for height, weight and BMI at each age, labelling measurements that fall outside this range 'extreme' and recommending that they be excluded from analyses.[18] According to these criteria, height z-scores outside the range of –6 to +6, weight z-scores outside the range of -6 to +5, and BMI z-scores outside the range of –5 to +5 are considered implausible. These cut-offs are well beyond those used to demarcate 'normal' height and weight.

Longitudinal data

The literature on identifying implausible variation within longitudinal anthropometric data is sparse. In the creation of the WHO Multicentre Growth Reference itself, the data were validated "on the basis of the range and consistency rules built into the data entry dictionary", with further checks to identify "measurements changing abnormally relative to the chronology of follow-up visits", and the examination of individual plots for "any questionable patterns".[8] In some cases, interviewers were sent back to re-measure the individual, but otherwise the protocols used for these processes are not described. Additional detail on these data consistency checks is provided in another article[9], but without explanation of the methods underlying these 'checks':

> *For anthropometry, the data entry system included built-in range and consistency checks that flagged measurements exceeding ±2 standard deviations of age- and sex-specific reference values for attained size. Flagged values were then checked for consistency between the two observers, consistency with other anthropometric variables measured on the same visit, consistency with previous measurements of the same child, and possible data entry errors.*

Many researchers have conducted longitudinal analyses using the height and weight data from the LSAC. However, the LSAC documentation does not describe any procedures used to assess the accuracy of anthropometric data. None of the identified articles published about the LSAC describe methods for evaluating the plausibility of intraindividual variation.[19–23] One of these articles stated that children with 'extreme BMI values (i.e. >40 kg/m^2)' were excluded from analyses; there was no justification for this cut-off point (or its relevance for children of different ages), and no description of efforts to identify 'extreme' intraindividual variation.[22]

Two studies were identified that described the assessment of intraindividual changes in height or weight[1,7]; however, neither provided justification for the selection of cut-off points (see Table 2). Given the absence of a comprehensive method in the literature, further advice was sought from the LSIC team, epidemiologists, paediatricians, nutritionists and endocrinologists to inform the development of an approach.

Approach for identifying likely inaccurate height and weight data in LSIC

An approach was developed to identify measurements in LSIC that were likely to be inaccurate. The DSS provided the height and weight data in raw form, and the plausibility of the data was assessed by (a) examining deviation from the WHO reference population by

Table 2. Studies describing the assessment of intraindividual changes in height or weight

Paper	Data	Approach
Noel and colleagues[1]	More than 20 million adults in the US Veterans Health Administration Corporate Data Warehouse	The authors describe the unfeasibility of examining individual trends to identify improbable data patterns in such a large study. Their solution was to assign cases with a small change in height or weight (1–2 cm for height or 10–100 lb for weight) 'within the realm of plausibility', moderate changes (2–10 cm for height or 100–1000 lb for weight) 'suspect' and larger changes (>10 cm for height or >1000 lb for weight) 'clearly implausible'. However, they do not provide a rationale or clinical basis for the choice of these classifications, or explicit description of how they processed these groups.
Harrist and Dai[7]	678 children aged 8–18 years from the Project HeartBeat Study	Although the group did not explicitly describe the use of z-scores, they used multilevel models to examine intraindividual variability away from the individual's 'trajectory', flagging points more than 3 standard deviations away from the subject-specific trajectory. The rationale for selecting the cut-off point of 3 standard deviations was not stated. Participants with flagged measurements were then examined in detail by the steering committee and were either corrected or set to missing. The criteria used to determine the 'appropriate corrective action' for these flagged data, however, are not stated.

calculating height, weight and BMI z-scores adjusted for age and sex, and (b) examining changes in individuals over time. Data were analysed using Stata version 12. The WHO Anthro and AnthroPlus macros for Stata were used to transform the raw age, height and weight data from LSIC into height, weight and BMI z-scores for analysis.[18,24]

If a height, weight or BMI z-score in LSIC fell outside the plausible range described by WHO[18], the measurement was excluded from analyses.[11] The prevalence of implausible data decreased across waves of the study (e.g. from 13% to 3% for BMI z-scores; see Table 3), consistent with RAOs' perceptions of improved accuracy.

Table 3. BMI z-scores recorded in each wave that were flagged as implausible, and those remaining after exclusion criteria were used

	Wave 1 (%)	Wave 2 (%)	Wave 3 (%)	Wave 4 (%)	Total (%)
BMI z-score recorded	1234	1371	1270	1233	5108
BMI z-score considered 'extreme' (excluded)	155 (12.56)	94 (6.86)	37 (2.91)	32 (2.60)	318 (6.23)
BMI z-score remaining after full exclusion criteria used	996 (80.71)	1207 (88.04)	1149 (90.47)	1170 (94.89)	4522 (88.53)

In addition to excluding these extreme values, a method was developed to distinguish the implausible within-child variability from the natural variability expected during childhood growth. Individuals were flagged if they were recorded as decreasing in height (in centimetres) between any two waves of the study, because it is physiologically impossible for children to lose height, except in cases of severe pathology.

Individuals were also flagged if they were recorded as having a 'significant' decrease in weight between any two waves. Because children can plausibly lose weight over time if they are sick or experience trauma, conservative exclusion criteria were employed to maintain the true biological variability represented in the data. Decreases in weight were considered 'significant' if the loss of weight (in kilograms) was associated with a decrease in weight z-score greater than 3. This cut-off point for intraindividual change – similar to that employed in Project HeartBeat[7] – was selected for LSIC because a change of this magnitude would represent a drastic shift in weight status, such as from obese to normal weight, or from normal weight to underweight, within a year.

Extreme increases in height or weight were not explicitly flagged, to allow for different growth trajectories for children, such as an early or late growth spurt, and because a reasonable definition for an implausible increase in height or weight could not be determined. However, because these data are longitudinal, an extreme increase in height or weight between waves was flagged for exclusion if the succeeding measurement decreased to the individual's earlier trajectory.

An algorithm was developed to identify the height and weight measurement(s) to be excluded from the sequence of measurements for flagged individuals. This protocol was based on individual z-score trajectories for height and weight. Given the observed tracking of height and weight status over time[25], the z-score trajectory with the least fluctuation between waves was considered the most plausible trajectory. Therefore, the data point(s) to be excluded were selected to minimise the change in an individual's z-score between successive waves. The criteria were (see Thurber[11] for more detail):

1. If there are only two data points recorded for an individual, it is not possible to infer which is more likely

to be in error based on comparison with a third data point; thus, exclude both.

2. If there are three or more data points recorded for an individual, use height-for-age or weight-for-age z-scores (in the case of decreases in height and decreases in weight, respectively) to determine which data point should be excluded based on consistency: remove the data point such that the sum of the differences in z-scores between successive waves is minimised.
3. If there are two non-consecutive sets of decreases in height or weight within the four data points (e.g. a decrease in height between wave 1 and wave 2, an increase in height between wave 2 and wave 3, and a decrease in height between wave 3 and wave 4), exclude all four data points because it is difficult to infer which data points are most likely to be in error.

As the exclusion method was based on z-scores, the validity of measurements for children missing height, weight or age could not be assessed (since z-scores could not be calculated). Thus, children missing data for any of these variables were not included in analyses. After the exclusion processes, the final sample included around 1000 BMI z-score measurements in each wave, representing 81–95% of all the BMI z-score measurements originally recorded.

Discussion

The processes outlined in this paper were designed to allow most of the dataset's original variability to be maintained, and to only exclude data points with a relatively high probability of representing measurement error. As a result, more data points fall at either extreme end of the BMI distribution than would be expected in a normal distribution. This effect persists through the fourth wave of the study, when the accuracy of data collection is presumed to have been improved. This likely represents the heterogeneity of weight status among the Australian Aboriginal and Torres Strait Islander population, which has been documented in other studies.[26–28]

From the original sample, anthropometric data for Torres Strait Islander (compared to Aboriginal) children and for children from areas with the highest levels of remoteness are under-represented. This should be considered when undertaking analyses of these data; however, given that LSIC data are not intended to be representative of the entire Australian Indigenous population, this should not discourage the use of these data, particularly for the conduct of internal comparisons or longitudinal analyses. The anthropometric data in LSIC constitute the largest available source of information about the longitudinal growth of a geographically diverse sample of Australian Aboriginal and Torres Strait Islander children.

The use of height, weight and BMI z-scores has associated limitations. The reference used for standardisation varies across studies; some researchers promote the use of references that are specific to factors such as ethnicity or country. However, research has shown that the variability in height and weight across countries or ethnicities is insignificant in comparison with the variability attributable to socioeconomic status, health and nutrition.[14] Indigenous Australian children have demonstrated a similar growth potential to non-Indigenous Australian children[29]; thus, a reference specific to Indigenous Australians is unnecessary. The use of the WHO reference has considerable support across settings, particularly for the Australian Indigenous population.[30]

Conclusion

This paper outlines an a priori approach for assessing the plausibility of anthropometric data in a longitudinal study, and identifying measurements that are likely to be errors. The LSIC team accepted this approach for identifying implausible data, resulting in the release, in 2012, of 'cleaned' data for public use for the first time.[31] This approach will allow children and their families to realise benefit from their participation in the measurement process. The approach documented here has been adapted by the DSS into an iterative program to enable the automatic evaluation of future waves of data; the fifth wave of data was cleaned using this program and released in April 2014 with Release 5.0.

Although it is not possible to directly evaluate the accuracy of these measurements, short of revisiting and re-measuring each queried individual, this method presents a probabilistic model to identify implausible measurements for exclusion. This protocol builds on the available literature and guidelines, and considers the clinical significance.

This protocol is not intended to displace the critical importance of accurate measurement techniques, but rather enables the improvement in accuracy of anthropometric data that have been collected. This protocol is systematic, practical and intended to be reproducible in other settings, including for the verification of large databases. Although Noel et al. state that "The massive volume of data that is typically available limits the capacity to develop algorithms to eliminate errors"[5], this paper presents an approach, based on WHO standards and the available evidence, to systematically eliminate less plausible measurements and trajectories in a dataset with more than 4000 measurements.

Acknowledgements

Footprints in Time – the Longitudinal Study of Indigenous Children (LSIC) would never have been possible without the support and trust of the Aboriginal and Torres Strait Islander families who opened their doors to the researchers and generously gave their time to talk openly

about their lives. Our gratitude goes to them, and to the leaders and elders of their communities who are active guardians of their people's wellbeing.

The authors acknowledge that LSIC is designed and managed under the guidance of the LSIC Steering Committee. The Committee, chaired by Professor Mick Dodson AM since 2003, has a majority of Indigenous members who have worked with the DSS to ensure research excellence in the study's design, community engagement, cultural, ethical and data access protocols, and analyses for publication. We would like to thank the LSIC Steering Committee and the RAOs for their support and assistance in this endeavour.

The authors also acknowledge the LSIC RAOs for their insight and assistance and the DSS team, in particular Laura Bennetts-Kneebone and Ana Sartbayeva, for their continued support and cooperation. The authors also acknowledge the assistance of Dr Phyll Dance, Dr Gillian Hall and Dr Martyn Kirk of the Australian National University, and Dr John Boulton of the universities of Sydney and Newcastle, for assistance in developing these methods.

This paper uses unit record data from LSIC, which has been initiated, funded and managed by the DSS. The findings and views reported in this paper, however, are those of the authors and should not be attributed to the DSS or the Indigenous people and their communities involved in the study. This work was supported by the Fulbright-Anne Wexler Master's Scholarship in Public Policy, the Australian National University and the National Centre for Epidemiology and Population Health. EB is supported by the National Health and Medical Research Council.

Competing interests

Work by EB on this research, and other research projects, is funded by an NHMRC Senior Research Fellowship.

References

1. Noel PH, Copeland LA, Perrin RA, Lancaster AE, Pugh MJ, Wang CP, et al. VHA Corporate Data Warehouse height and weight data: opportunities and challenges for health services research. J Rehabil Res Dev. 2010;47:739–50.
2. Vioque J, Ramos JM, Navarrete-Muñoz EM, García-de-la-Hera M. A bibliometric study of scientific literature on obesity research in PubMed (1988–2007). Obes Rev. 2010;11(8):603–11.
3. Stevens J, Truesdale KP, McClain JE, Cai J. The definition of weight maintenance. Int J Obes. 2006;30(3):391–9.
4. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. Br J Nutr. 1999;82(03):165–77.
5. Das SR, Kinsinger LS, Yancy WS Jr, Wang A, Ciesco E, Burdick M, et al. Obesity prevalence among veterans at Veterans Affairs medical facilities. Am J Prev Med. 2005 Apr;28(3):291–4.
6. Noel PH, Copeland LA, Pugh MJ, Kahwati L, Tsevat J, Nelson K, et al. Obesity diagnosis and care practices in the Veterans Health Administration. J Gen Intern Med. 2010;25(6):510–6.
7. Harrist RB, Dai S. Analytic methods in Project HeartBeat! Am J Prev Med. 2009;37(1):S17–S24.
8. Onyango AW, Pinol AJ, de Onis M. Managing data for a multicountry longitudinal study: experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull. 2004 Mar;25(1):46S–52S.
9. De Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull. 2004 Mar;25(1):S27–36.
10. Thurber KA, Banks E, Banwell C. Cohort profile: footprints in time, the Australian longitudinal study of Indigenous children. Int J Epidemiol. 2014. pii: dyu122. Epub 2014 Jul 9.
11. Thurber KA. Analyses of anthropometric data in the longitudinal study of Indigenous children and methodological implications [masters thesis]. [Canberra]: Australian National University; 2012. Available from: http://hdl.handle.net/1885/9892
12. The Longitudinal Study of Indigenous Children. Data user guide release 4.0. Canberra: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs; 2013.
13. World Health Organization. WHO child growth standards methods and development: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. Geneva: World Health Organization; 2006.
14. World Health Organization. Physical status: the use and interpretation of anthropometry. WHO technical report series. Geneva: World Health Organization; 1995.
15. Gorstein J, Sullivan K, Yip R, de Onis M, Trowbridge F, Fajans P, et al. Issues in the assessment of nutritional status using anthropometry. Bull World Health Organ. 1994;72(2):273–83.
16. De Onis M, Lobstein T. Defining obesity risk status in the general childhood population: which cut-offs should we use? Int J Pediatr Obes. 2010;5(6):458–60.
17. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. BMJ. 2007;335(7612):194.

18. Blossner M, Siyam A, Borghi E, Onyango A, de Onis M. WHO AnthroPlus for personal computers manual: software for assessing growth of the world's children and adolescents. Geneva: World Health Organization, Department of Nutrition for Health and Development; 2009.

19. Jansen PW, Mensah FK, Nicholson JM, Wake M. Family and neighbourhood socioeconomic inequalities in childhood trajectories of BMI and overweight: Longitudinal Study of Australian Children. PLoS One. 2013;8(7):e69676.

20. Jansen PW, Mensah F, Clifford S, Nicholson JM, Wake M. Bidirectional associations between overweight and health-related quality of life from 4–11 years: longitudinal study of Australian children. Int J Obes. 2013;37(10):1307–13.

21. Magee CA, Caputi P, Iverson DC. Patterns of health behaviours predict obesity in Australian children. J Paediatr Child Health. 2013 Apr;49(4):291–6.

22. Magee CA, Caputi P, Iverson DC. Identification of distinct body mass index trajectories in Australian children. Pediatr Obes. 2013;8(3):189–98.

23. Wake M, Maguire B. Longitudinal study of Australian children. Annual statistical report 2011. Melbourne: Australian Institute of Family Studies; 2012.

24. World Health Organization. WHO Anthro (version 3.2.2 January 2011) and macros. Geneva: WHO; 2012 [cited 2012 Oct 1]. Available from: www.who.int/childgrowth/software/en/

25. Singh AS, Mulder C, Twisk JWR, Van Mechelen W, Chinapaw MJM. Tracking of childhood overweight into adulthood: a systematic review of the literature. Obes Rev. 2008;9(5):474–88.

26. Wang Z, Hoy W, McDonald S. Body mass index in Aboriginal Australians in remote communities. Aust N Z J Public Health. 2000;24(6):570–9.

27. Cunningham J, Mackerras D. Overweight and obesity: Indigenous Australians 1994. Occasional paper. Canberra: Australian Bureau of Statistics; 1998.

28. Schultz R. Prevalences of overweight and obesity among children in remote Aboriginal communities in central Australia. Rural Remote Health. 2012;12(1872).

29. Gracey M. Nutrition of Australian Aboriginal infants and children. J Paediatr Child Health. 1991;27:259–71.

30. Department of Health and Community Services. Child growth charts in the Northern Territory: discussion paper. Darwin: Northern Territory Government; 2008.

31. The Longitudinal Study of Indigenous Children. Data user guide release 3.1. Canberra: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs; 2012.