

## ***Growing Up in Australia: The Longitudinal Study of Australian Children***

### **Height differences**

In the leave-behind questionnaires for both parents at Wave 1 and Wave 2, the parents were asked to report their height and weight so their body mass index (BMI) could be calculated. In cleaning Wave 2 data it was discovered that there was a large number of discrepancies between the values reported by the same people at Wave 1 and at Wave 2. In fact, only 50 per cent of respondents reported a value that was within 1 per cent of their Wave 1 value.

Further investigation failed to find any explanation other than respondent error for the vast majority of these cases. In order that data analysts could assume that any observed changes in BMI were due to changes in reports of weight rather than height, it was decided to impute the value of height to be the average of the two reported values.

At Wave 3, the question on the Parent 1's height was asked of all new Parent 1's and those that had not returned a self-complete form at Wave 2, plus a handful of cases where Parent 1 had swapped places with Parent 2. However, for Parent 2 the height data was still collected by self-complete form, so sequencing cases around the question was not an option. Hence, for many<sup>1</sup>, there are now 3 points of data collection.

When the study child's height is measured as part of the interview process, a third measurement is taken if the first two disagree by more than 0.5cm. If this is the case the estimate of the child's height is considered to be the average of the two that correspond the most closely. This method of estimation means that the least reliable estimate has no effect on the result. It is suggested that in cases with three data points for a parents height, the 'clean' result provided on the data file could similarly be the average of the closest two responses. As is done currently, the values of parental height for each wave prior to this cleaning will remain on the data file if analysts wish to use their own approach.

Figure 1 shows the discrepancy between the two values used to create the 'clean' result for those parents with two data points versus those with three. Those with three data points had two that agreed in 77 per cent of cases. Those with two data points had agreement in only 42 per cent of cases. It should be noted, however, that at Wave 2, 45 per cent of cases had agreement between the two data points, so there is some evidence that those who were more likely to return self-complete questionnaires were more likely to give accurate data.

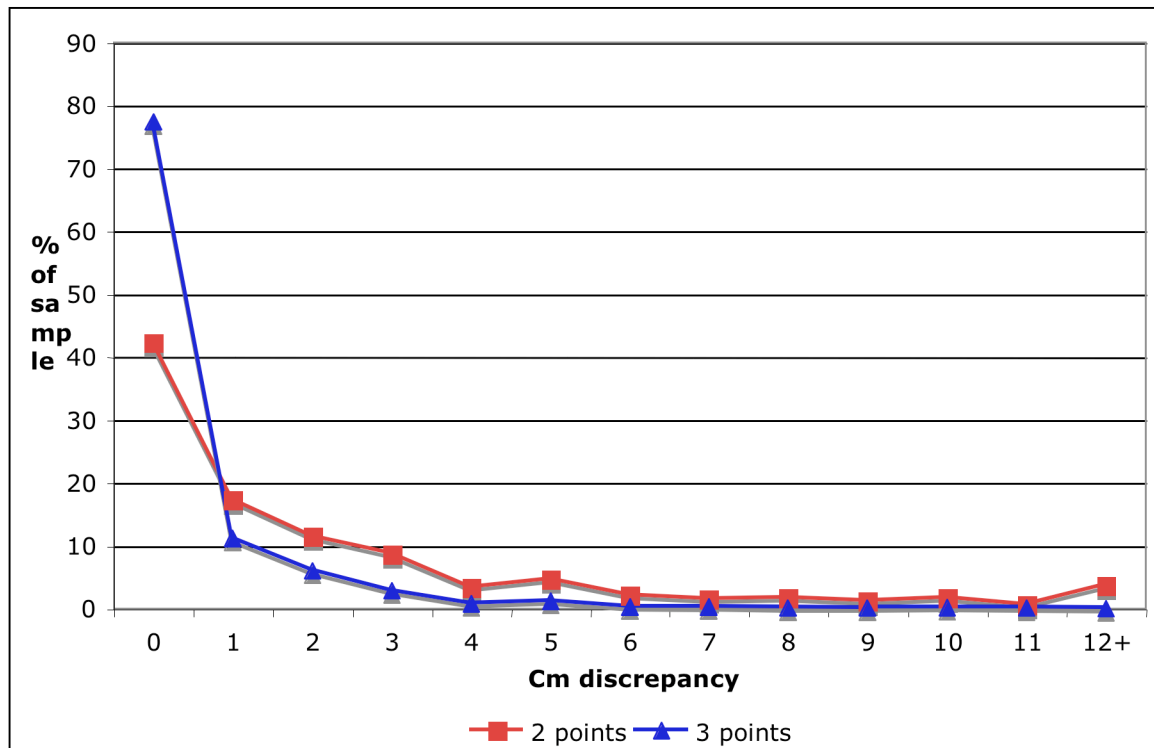
It was decided on inspection of Figure 1 that any case with more than 10cm discrepancy between the two closest values should be considered unreliable and therefore should be set to missing. This would affect 4 per cent of Wave 3 parents with two data points and less than 0.1 per cent of cases with three data points.

---

<sup>1</sup> 72 per cent of those who return Parent 2 questionnaires at Wave 3 have data from all 3 waves. Of all Wave 3 Parent 2s, 11% had no height data, 16% had one data point, 22% had two data points and 51% had 3 data points.

*Outcome*

- This problem with the height data was presented to the February 2009 data Expert Reference Group Meeting and the group decided that if the differences are less than 10cm then average all three, otherwise average the closest two. Consequently this is how the height data has been adjusted.*



**Figure 1. Centimetre discrepancy in two closest data points for those with three versus two data points on parental height for Wave 3 respondents**