

**Title: BMI peak in infancy as a predictor for later BMI in the Uppsala Family Study**

Running title: BMI peak in infancy as a predictor for later BMI

**Authors: Richard J Silverwood<sup>1</sup>, Bianca L De Stavola<sup>1</sup>, Tim J Cole<sup>2</sup>, David A Leon<sup>1</sup>**

<sup>1</sup>Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT

<sup>2</sup>MRC Centre of Epidemiology for Child Health, UCL Institute of Child Health, London, WC1N 3JH

Correspondence to:

Richard Silverwood

Department of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

Keppel Street

London

WC1E 7HT

Richard.Silverwood@lshtm.ac.uk

0207 927 2525

## Abstract

### Background

The development of obesity through childhood, often characterised using body mass index (BMI), has received much recent interest due to the rapidly increasing levels of obesity worldwide. However, the extent to which BMI trajectory in the first year of life (the BMI ‘peak’ in particular) is associated with BMI in later childhood has received little attention.

### Subjects

The Uppsala Family Study includes 602 families, comprised of mother, father and two consecutive singleton offspring, both of whom were delivered at the Uppsala Academic Hospital, Sweden, between 1987 and 1995. Children's postnatal growth data, including serial measurements of height and weight (from which BMI was calculated), were obtained from health records. All children had a physical examination when they were aged between 5 and 13 years, at which height and weight were again recorded and used to calculate age- and sex-adjusted BMI z-scores.

### Methods

Subject-specific growth curves were fitted to the infant BMI data using penalised splines with random coefficients, and from these the location of the BMI peak for each subject was estimated. A multilevel modelling approach was used to assess the relationships between the BMI peak and BMI z-score in later childhood.

## Results

The BMI peak occurred, on average, slightly later in females, with a higher BMI peak in males. Considered separately, both age and BMI at BMI peak were positively associated with later BMI z-score. Considered jointly, both dimensions of BMI peak retained their positive  
5 associations.

## Conclusions

The growth trajectory associated with higher childhood BMI appears to include a later and/or  
higher BMI peak in infancy.

10

## Keywords

Infancy, Obesity, Childhood obesity, Body mass index, Growth

## Introduction

Obesity in childhood is increasing rapidly worldwide [1], bringing with it many adverse health and social consequences. The acknowledged pattern of tracking of obesity from childhood to adulthood [2] means that even if overweight children avoid health problems in their youth, they have an increased likelihood of being overweight, and thus encountering the associated increases in type II diabetes, cardiovascular disease risk factors, and respiratory and psychosocial problems [3, 4] in adulthood.

Obesity is often considered in terms of body mass index (BMI), calculated as  $\text{weight}/\text{height}^2$ , due to the ease with which measurements can be made. For a typical individual, BMI increases from birth until around age 9 months where it reaches a maximum, the ‘BMI peak’. It then decreases, reaching a nadir (the ‘BMI rebound’, also referred to as the ‘adiposity rebound’) around age 6 years, before increasing once more. Whilst from Rolland-Cachera et al [5] onwards [6] earlier BMI rebound has been consistently shown to be associated with later obesity, little research has been conducted into possible associations between the earlier BMI peak and later obesity.

Although not considering the BMI peak explicitly, many studies have investigated the relationships between size and/or growth in infancy and later obesity. A systematic review by Baird et al [7] identified 18 studies concerning infant size, 11 of which found that infants who were heavier during infancy or were defined as obese were more likely to develop obesity in childhood, adolescence or adulthood. Several reviews, ranging in size from 10 to 21 studies, have focused on infant growth as an exposure [7, 8, 9]. The majority of studies in each review found more rapid growth in infancy to be associated with a greater risk of subsequent obesity.

However, the definition of ‘infant growth’ is somewhat inconsistent, with increases in absolute weight, weight for age z-score or weight for height z-score over a fixed period after birth all utilised. Ong and Loos [9] attempted to standardise the results, concluding that there is strong evidence for a true association between rapid infancy weight gain and subsequent risk of obesity in later life. A more recent assessment of the nature and strength of the evidence for this association found it to be supported by observational studies but not by the relatively small randomised trials [10]. Further, the statistical methods used to estimate individual measures of growth are dictated by the available observations and do not fully exploit the data. We have the opportunity to investigate the relationship between the timing and extent of the BMI peak and BMI z-score in later childhood in the Uppsala Family Study (UFS), a dataset of 1204 subjects from Uppsala, Sweden. This dataset is unusual in having a large number of serial measurements of height and weight from birth into childhood, which we use to estimate age and BMI at BMI peak using penalised splines with random coefficients.

15

## **Subjects**

20 The UFS population comprised families with two full-siblings and their biological mother and father [11]. All families with at least two consecutive singleton children delivered at term and within 36 months of each other at the Uppsala Academic Hospital, Uppsala, Sweden, between 1987 and 1995 were potentially eligible for the study. Children also had to share the same biological father and families had to live within Uppsala county at the time of the study, with  
25 both parents of Nordic origin. If there were more than two children in the same family

fulfilling these criteria then the oldest two siblings were chosen. By linkage between the Swedish Medical Birth Registry and the current population register, 5226 women and their 10,452 offspring were identified as fulfilling these criteria and hence comprised the sampling frame for the study.

5

The initial focus of the data collection was to study maternal and early life effects on blood pressure and cardiovascular disease [12]. To increase statistical efficiency only families where the siblings were either both in the top or bottom quarter of the sex-specific birthweight distribution ('concordant high birthweight' or 'concordant low birthweight') or the sex-adjusted difference in birthweight between them was 0.4 kg or more ('discordant birthweight') were invited to participate. However, in the context of this paper the study design provides no benefits in terms of efficiency. A total of 1967 families were invited to take part in the study, with 71% responding, leading to the eventual recruitment of 602 families (31% of those eligible). Siblings could be same-sex pairs or discordant for sex.

15

Children's birth data were obtained from mothers' obstetric records through the Swedish Medical Birth Registry, and their postnatal growth data, including serial measurements of height and weight, were obtained from health records kept by Child Health Centres or schools. The measurements of height and weight were used to calculate BMI.

20

All 1204 children in the UFS had a physical examination between May 2000 and November 2001 when they were aged between 5 and 13 years, at which many measurements were recorded. Of these, only height and weight were utilised in the present analysis. Height was measured with a wall-fixed stadiometer to an accuracy of 0.1 cm with subjects walking around the room between measurements, and weight was measured with the subject wearing

25

underwear to an accuracy of 0.1 kg using electronic scales. Height and weight measurements were taken three times and the mean value used. From the observed height and weight values BMI was again calculated, from which age- and sex-adjusted BMI z-scores were obtained using the Swedish population reference values [13].

5

Preliminary exploratory analyses (not shown) estimated the BMI peak to occur at an age of between 6 months and 1 year in the majority of individuals. Thus only BMI measurements between birth and age 3 years were included in the BMI growth curve modelling. Forty of the initial 1204 individuals (3.3%) had no BMI observations within this range so were excluded.

10 A further 2 subjects (0.2%) with fewer than 3 datapoints, the minimum number deemed necessary, were excluded, leaving 1162. The proportion of excluded subjects was similar across sexes, sibling types and birthweight groups..

15

## **Methods**

The analysis may be considered as a two stage process. First, infant BMI data were used to construct subject-specific BMI growth curves from which the BMI peak was identified. Then  
20 assessment was made of the relationships between features of the BMI peak and later BMI z-score.

Subject-specific growth curves were fitted to the infant BMI data using penalised splines with random coefficients. The formulation of this modelling approach is described in the Annex  
25 and in greater detail elsewhere [14].

Briefly, penalised splines are a means of modelling the relationship between two (or more) variables without the imposition of a rigid parametric form (for example linear or higher order polynomial). They are formed by the piecewise connection of polynomial curves between a series of ‘knots’ subject to certain penalties to encourage smoothness. Mixed models [15] may be fitted to datasets comprising repeated measures across multiple individuals in order to obtain subject-specific growth curves [16]. As penalised splines can be handled within the mixed model framework they can also be extended in a similar manner. This fusion between parametric mixed modelling and smoothing is referred to as ‘semiparametric mixed modelling’ [14]. The mixed model representation means that relatively complex penalised spline models can be implemented using standard statistical software.

Using this approach, only the number of knots and the knot locations need be specified. We used 12 knots, a similar number to that used in comparable applications elsewhere [17], though perhaps towards the upper limit of what is necessary. A simple approach to selecting the knot locations is on the basis of the quantiles of the unique observed values of the independent variable [14], so knots were placed at the  $\left(\frac{1}{13}\right)^{th}, \dots, \left(\frac{12}{13}\right)^{th}$  quantiles of the unique ages.

Acknowledged differences in childhood BMI growth between males and females [18] suggest that different underlying growth trajectories should be used for each sex. Moreover, different growth patterns may be expected for individuals with different birthweights [19]. For these reasons six separate models were fitted (concordant low birthweight males, concordant high birthweight males, discordant birthweight males, concordant low birthweight females, concordant high birthweight females and discordant birthweight females).

As the BMI peak is a turning point in the BMI curve, the age at BMI peak for an individual can be estimated via differentiation (with respect to age) of their fitted BMI growth curve. BMI at BMI peak is then the BMI value corresponding to this age according to the BMI growth curve. Whilst this simple approach to identifying the BMI peak worked well for most individuals, some subjects had fitted BMI growth curves with local non-BMI peak maxima. It was thus additionally required that the first derivative of the BMI curve must be positive 3 months before and negative 3 months after a maximum for it to be considered the BMI peak.

In the second stage, two features of the BMI peak, the corresponding age and BMI, were related to BMI z-score in later childhood. To account for the dependencies caused by the inclusion of sibling pairs in the UFS, a multilevel (mixed) modelling approach [15] was used, with family-specific random intercepts. Sex, birthweight and age at physical examination were included, with interactions between BMI peak and each of these variables also considered.

## Results

Table 1 summarises the distributions of several variables assessed at birth or at the physical examination in later childhood in the 1162 subjects with at least 3 BMI observations between birth and age 3 years. Mean weight and length at birth were slightly higher in males than females, though BMI z-score at physical examination was higher in females. Average age at physical examination was approximately 10 years in both sexes, though there was much variability. Equivalent summaries for the 42 excluded subjects with fewer than 3 BMI

observations between age 0 and 3 years (results not shown) showed no substantial differences from the values in Table 1.

[Table 1 here]

5

The number of BMI observations for each subject varied greatly between the 1162 individuals (median 13, interquartile range 6, range 3-30). Additionally, the distribution of measurement ages was far from uniform between birth and age 3 years - over 50% of data corresponded to ages less than 6 months, and data were markedly more sparse for ages greater than 1.5 years.

10

#### BMI growth curves

The population average curves predicted for each birthweight group are plotted separately for males (upper plot) and females (lower plot) in Fig. 1. The curves took a similar shape in each birthweight group, indicating that the shape of the BMI trajectory in infancy is largely independent of birthweight, though there were clear vertical displacements of the curves. The curves for males were as would be expected, with the concordant low birthweight subgroup having a lower trajectory right across the range of ages examined, the concordant high birthweight subgroup having a higher trajectory, and the discordant birthweight subgroup being between the two. The observed trends in females were very similar, although here the trajectory of the discordant birthweight subgroup much more closely matched that of the concordant high birthweight subgroup. For both males and females the ages at which the population average curves reached the maximum differed somewhat between the birthweight groups. Whilst some of the curves appear to be plateauing at later ages rather than displaying the anticipated decrease in BMI towards the BMI rebound it should be remembered that, as

25

well as there being relatively few datapoints at these ages, these are population average curves, so individual subject-specific curves may be either increasing or decreasing at any given age.

5 [Fig. 1 here]

The combination of these population average curves and the estimated subject-specific deviations from them gives the overall fitted subject-specific BMI curves. These are presented in Fig. 2 for several males selected in a systematic manner, as described in the figure caption.

10 This collection of plots provides examples of each subgroup model for varying levels of infant BMI. Whilst the subject-specific curves all took the same general shape as the population average curves, the inclusion of the random coefficients allowed the subject-specific curves to, on the whole, provide good fits to the data. The number of data points per subject was variable, and for individuals where data were more sparse fitted curves drew  
15 information from others. Although most of the curves in Fig. 2 showed obvious maxima, some were flatter than others. Equivalent plots for females (not shown) indicated similar patterns.

[Fig. 2 here]

20

Estimated age and BMI at BMI peak

Table 2 summarises the distributions of age and BMI at BMI peak, along with the number and percentage of subjects with identified BMI peak, by sex and birthweight group. The percentage of subjects with a successfully identified BMI peak was generally high, though

some differences between the birthweight groups were evident, with identifiability being greater for smaller babies, and also in females.

[Table 2 here]

5

The BMI peak occurred later in concordant high birthweight males and in concordant low birthweight females than in the other birthweight groups (both  $P < 0.001$  by t-test). Overall, BMI peak occurred later in females than males ( $P < 0.001$  by t-test). The median age at BMI peak was generally lower than the mean, suggesting a skewed distribution.

10

Average BMI at BMI peak was highest in concordant high birthweight subjects and lowest in concordant low birthweight subjects in both sexes, corresponding to the population average curves seen in Fig. 1. As a later BMI peak thus corresponded to the highest BMI at BMI peak in males (concordant high birthweight subjects) but to the lowest BMI at BMI peak in females (concordant low birthweight subjects) there was some evidence that the relationship between age and BMI at BMI peak may differ by sex and birthweight group. Generally, BMI at BMI peak was greater in males ( $P < 0.001$  by t-test). Mean and median were very similar in each group, indicating a more symmetric distribution.

15

20

The main reason for subjects not having an identified BMI peak was that their BMI observations continued to increase over the first few years of life. Comparing subjects with and without an identified BMI peak suggested that both males and females with no identified BMI peak generally had slightly greater weight and length at birth (results not shown).

25

Is BMI peak location related to BMI z-score in later childhood?

Table 3 details the separate effect of age and BMI at BMI peak on BMI z-score at physical examination estimated accounting for the within-family correlations. There was no strong evidence of interaction between sex and either dimension of the BMI peak, thus sex was included as a confounder, as were birthweight and age at physical examination. In particular, with reference to the observation that the relationship between age and BMI at BMI peak may differ by sex and birthweight group, there was no clear evidence of a 3-way interaction between age at BMI peak, sex and birthweight ( $P=0.09$ ). Both a delayed age at BMI peak and an increased BMI at BMI peak were associated with a positive and highly statistically significant increase in BMI z-score in later childhood. Birthweight was also strongly positively associated with BMI z-score. Further, for a given BMI at BMI peak and birthweight, females had a significantly higher BMI z-score at examination than males.

[Table 3 here]

15

Table 4 details the estimated joint effects of age and BMI at BMI peak, adjusted for sex, birthweight and age at physical examination. There was weak evidence ( $P=0.04$ ) of an interaction between age and BMI at BMI peak, suggesting that BMI at BMI peak may be less informative in those who exhibit a later BMI peak, but this is not considered further here. There were significant associations between both age and BMI at BMI peak and BMI z-score in childhood, even after mutual adjustment, although the evidence for the BMI at BMI peak association was somewhat stronger. There also remained a strong, positive relationship with birthweight. In this model, for a given birthweight and age and BMI at BMI peak, females were expected to have a higher BMI z-score at examination.

25

[Table 4 here]

One way to consider the model in Table 4 is to calculate predicted BMI z-scores for different combinations of the two exposures of interest. By holding the values of the other explanatory variables constant it is possible to examine predicted BMI z-scores for different combinations of age and BMI at BMI peak through use of a contour plot (Fig. 3, for males with mean birthweight and mean age at physical examination). The region of highest predicted BMI z-score in later childhood corresponded to a late BMI peak and a high BMI at BMI peak. The lowest predicted BMI z-scores corresponded to early BMI peak and a low BMI at BMI peak. It is clear from the plot that BMI as opposed to age at BMI peak exerted the greater influence.

[Fig. 3 here]

Equivalent contour plots using different combinations of values of sex, birthweight and age at physical examination would lead to the same interpretation. As these variables entered the model additively they would only have changed predicted BMI z-score in later childhood by a constant value, leaving the shape of the contour plot in Fig. 3 unchanged.

As birthweight is known to affect growth trajectories [19], the study design could result in the UFS being unrepresentative both in terms of the BMI peak locations and later BMI z-scores as well as, potentially, the relationship between the two. However, birthweight group was seen to have little effect on the shape of the BMI trajectory through infancy and birthweight was included as an explanatory variable in the second stage models, meaning any additive effect on BMI z-score should be adjusted for. Additionally, birthweight was not found to modify the relationships between the BMI peak and later BMI z-scores in any of the models.

We explored using models including both birthweight itself and indicator variables for birthweight group (results not shown). After adjustment for birthweight, birthweight group had no further effect on BMI z-score in any of the models, indicating that the effects of the study design were adequately accounted for.

5

We also calculated the residuals between each subject's observed age and BMI at BMI peak values and the expected values given their sex and birthweight. The absolute between-sibling pair differences in these residuals were then calculated and compared between concordant birthweight siblings (both concordant low birthweight and concordant high birthweight groups combined) and discordant birthweight siblings. No differences were observed in either age ( $P=0.4$ ) or BMI ( $P=0.9$ ) at BMI peak. These results suggest that once birthweight has been accounted for the greater heterogeneity in size in the discordant birthweight siblings has little further effect on their growth patterns. We thus believe that the unusual study design should not be allowed to detract unnecessarily from the results observed.

15

## **Discussion**

The initial peak in BMI at around the age of 6 months to 1 year was a readily identifiable feature of the BMI growth curve in the vast majority of subjects in the UFS. On average, the BMI peak occurred slightly later in females, with a higher corresponding BMI in males. Both higher BMI at BMI peak and later BMI peak tended to result in relatively higher BMI in later childhood. It is the first time that these associations have been reported.

Subjects did not contribute to the analysis for two reasons: either they had fewer than 3 BMI observations over the relevant ages so were excluded from the start, or it was not possible to identify the BMI peak from their BMI growth curve. Only 3.5% subjects had fewer than 3 BMI observations between birth and age 3 years, so their exclusion is unlikely to lead to substantial bias. However, although the BMI peak was identified in the majority of individuals considered, for 13% this was not the case. The observed BMI values for these subjects did not provide any evidence of a BMI peak, usually because BMI continued to increase throughout infancy. The analyses were effectively restricted to those subjects who exhibited a BMI peak. However, comparisons of those with/without a BMI peak showed no evidence of differences in BMI z-score in later childhood ( $P=0.3$ ).

We used penalised spline models with random coefficients to model BMI growth through infancy. For individuals with few observations the approach ‘borrows’ information from other subjects and fits a subject-specific curve closer to the relevant population average curve. This was an effective use of the available data which allowed us to identify the two dimensions of the BMI peak that we were interested in.

The study design of the UFS included the selection of sibling pairs based on their relative birthweights. We accounted for the sibling pairs by fitting random effects models which allowed subjects to be more similar to their sibling than to other members of the dataset. However, if all subjects were treated as independent using standard linear regression, the estimated model coefficients would change very little and the conclusions reached would be identical (results not shown). Additionally, the physical examinations at which the outcome was observed occurred across a wide range of ages (5-13 years). To account for this the models relating BMI z-score to dimensions of the BMI peak were adjusted for age at physical

examination, and age at physical examination was not found to modify the relationships between age and/or BMI at BMI peak and BMI z-score.

The inclusion criteria for the UFS provided a contemporary, healthy sampling frame, which was likely to be representative of the wider Uppsala population. However, as participation rates were not particularly high [12], subjects in the UFS may potentially not have been fully representative of those within the sampling frame.

Whilst no previous studies have explicitly investigated the effect of the location of the BMI peak on later levels of BMI, many studies have investigated the relationships between size and/or growth in infancy and later obesity [7, 8, 9, 10]. The majority of studies have quantified infant size and (particularly) growth in terms of weight, which we have not modelled, but some have used BMI. Although it is not possible to directly compare our results with those obtained previously as we did not estimate BMI at a given age or BMI gain over a given period for each subject, approximate comparisons can be made.

By categorising the 10% of subjects with the greatest (sex-specific) BMI at BMI peak as being ‘obese at BMI peak’ and the 10% of subjects with the highest BMI z-score in later childhood as being ‘obese in later childhood’ we found obesity at BMI peak to lead to an odds ratio (OR) of 5.20 (95% CI 2.22-12.21) for obesity in later childhood (adjusting for sex and birthweight and accounting for the within-family correlations). Whilst this is an admittedly crude comparison, this OR is greater than those reported previously when relating obesity in infancy to later obesity [7], perhaps indicating a greater predictive capability when the exposure is assessed at the BMI peak rather than at a fixed age in infancy. Further adjustment for age at BMI peak negligibly attenuated the OR.

We can also make an approximate comparison to previous results concerning infant growth by calculating the infant BMI gain rate (the average monthly increase in BMI between birth and BMI peak). By categorising the 10% of subjects with the greatest (sex-specific) infant BMI gain rate as having ‘rapid infant BMI gain’ we found rapid infant BMI gain to lead to an OR of 4.03 (95% CI 1.74-9.33) for obesity in later childhood. Although no previous studies focus on BMI gain in infancy, this OR is of comparable magnitude to many of those concerning infant weight gain [7, 8, 9].

If rapid infant BMI gain is associated with greater BMI and age at BMI peak then the relationship between rapid infant BMI gain and later BMI z-score could provide an explanation for the observed relationship between the two aspects of the BMI peak and later BMI z-score. However, whilst we found more rapid infant BMI gain to be positively associated with BMI at BMI peak, we found it to be negatively associated with age at BMI peak (both  $P < 0.001$ ). Thus the relationship between infant BMI gain and later BMI z-score cannot fully explain our results.

The observed positive relationship between BMI at BMI peak and later BMI could plausibly be explained by BMI tracking [2], which has been shown to occur from infancy to middle childhood [20]. If a subject has a high BMI relative to their peers at their respective BMI peaks then, even though these BMI values are not observed at the same age, it seems reasonable that they are likely to have a higher BMI several years later.

It is thus the positive relationship between age at BMI peak and later BMI which is perhaps the more interesting, particularly as this was in the opposite direction to that widely

acknowledged between age at BMI rebound and later BMI. This means that higher later BMI is associated with both those who are less well developed (in terms of BMI) around age 1 year (those having a later BMI peak) and those who are more well developed around age 6 years (those having an earlier BMI rebound), which is perhaps surprising. This leads to further questions regarding the relationships between these two features of the BMI growth curve and later BMI. For example, is it the same individuals who have both later BMI peak and earlier BMI rebound, leading to increased later BMI? Age at BMI peak and age at BMI rebound are both measures of development at that point, with regards to the BMI growth curve at least, and thus an inverse relationship between them would seem unlikely. Are there then disparate subgroups who have either a later BMI peak or an earlier BMI rebound and then proceed to increased later BMI? To answer these questions it is essential to have a dataset in which both the BMI peak and the BMI rebound can be identified for each individual. Unfortunately the present dataset does not afford the opportunity for this. Addressing these questions could provide valuable insights into BMI development through childhood.

Indeed, were our dataset appended with sufficient data to extend our models to include the BMI rebound, the plateauing seen at older ages in Fig. 2 for some subjects would likely change to the expected decrease. Whilst this could change the locations of the BMI peaks from those currently identified, we would expect any difference to be negligible.

As when considering any feature of childhood growth, a key question is whether the location of the BMI peak is a causal factor for later BMI itself or whether both the location of the BMI peak and later BMI are merely expressions of some genetic predisposition [21]. If it is causal, then can it be manipulated? Whilst the level of BMI for an infant, and thus their BMI at BMI peak, could be manipulated by changes in dietary intake, it remains unclear whether this

would have any effect on the timing of the BMI peak. It is also unclear whether age at BMI peak could be similarly manipulated. Furthermore, the imposition of dietary limitations on infants may be considered undesirable. This is a further area where additional research could prove fruitful.

5

Although the associations found in the present analysis between the BMI peak in infancy and later BMI z-score are of great interest, the UFS is a relatively small dataset meaning that replication in further datasets is necessary to confirm the veracity of the findings. Doing so may prove valuable in improving understanding of BMI development through childhood.

10

## **Acknowledgements**

We would like to thank Ilona Koupil, Torsten Tuvemo and Ann-Christine Synvanen for agreeing to providing access to these data and Rawya Mohsen for organising the data. The

5 original data collection was funded by the UK Medical Research Council.

## Annex

Consider initially the fitting of a growth curve for a single subject. Let  $y_j$  be the (log transformed) BMI for this subject at age  $x_j$ ,  $j = 1, \dots, n$ . Let  $\kappa_1, \dots, \kappa_K$  be a set of  $K$  distinct knots in the range of  $x_j$  and define

$$x_+ = \max(0, x).$$

Then the spline model of degree  $p$  with  $K$  knots at  $\kappa_1, \dots, \kappa_K$  is defined as

$$y_j = \beta_0 + \beta_1 x_j + \dots + \beta_p x_j^p + \sum_{k=1}^K u_k (x_j - \kappa_k)_+^p + \varepsilon_j, \quad (1)$$

where  $\beta_1, \dots, \beta_p$  and  $u_1, \dots, u_K$  are to be estimated and  $\varepsilon_j \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ . As unconstrained fitting of  $u_1, \dots, u_K$  will result in a ‘wiggly’ fit [14], a constraint such as  $\sum_{k=1}^K u_k^2 < C$  for some constant  $C$  may be imposed. The resulting model is referred to as a penalised spline model.

It can be shown [14], using the principle of ‘best linear unbiased prediction’ (BLUP) [22], that the penalised spline model (1) can be represented as a mixed model with  $u_k \stackrel{iid}{\sim} N(0, \sigma_u^2)$ . The model can then be considered in the general linear mixed model form,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_1 & \cdots & x_1^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^p \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} (x_1 - \kappa_1)_+^p & \cdots & (x_1 - \kappa_K)_+^p \\ \vdots & \ddots & \vdots \\ (x_n - \kappa_1)_+^p & \cdots & (x_n - \kappa_K)_+^p \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix}, \mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_K \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \text{ and } \text{Cov} \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} \sigma_u^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I} \end{pmatrix}.$$

5 Subject-specific growth curves are obtained via the inclusion of subject-specific random (spline) parameters which model the deviation of a given individual's curve from the population average (spline) curve. Now consider all subjects in the dataset so that  $y_{ij}$  is the (log transformed) BMI for subject  $i$ ,  $i=1, \dots, m$ , at age  $x_{ij}$ ,  $j=1, \dots, n_i$ . Then the penalised spline model of degree  $p$  can be extended to give

10

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \dots + \beta_p x_{ij}^p + \sum_{k=1}^K u_k (x_{ij} - \kappa_k)_+^p + a_{i0} + a_{i1} x_{ij} + \dots + a_{ip} x_{ij}^p + \sum_{k=1}^K v_{ik} (x_{ij} - \kappa_k)_+^p + \varepsilon_{ij} \quad (2)$$

where  $u_k \stackrel{iid}{\sim} N(0, \sigma_u^2)$ ,  $(a_{i0}, \dots, a_{ip})^T \stackrel{iid}{\sim} N(0, \boldsymbol{\Sigma})$ , where  $\boldsymbol{\Sigma}$  is an unstructured  $(p+1) \times (p+1)$

covariance matrix,  $v_{ik} \stackrel{iid}{\sim} N(0, \sigma_v^2)$ ,  $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ , and all terms are independent of one

15 another apart from  $a_{i0}, \dots, a_{ip}$  for a given  $i$ . The present analysis uses cubic penalised spline models, with both cubic population average curves and cubic subject-specific deviations from these (i.e. (2) with  $p=3$ ), to model BMI growth through infancy.

## References

- 1 World Health Organisation. *Obesity: preventing and managing the global epidemic*.  
World Health Organisation, Geneva, 2004.
- 5 2 Serdula MK, Ivery D, Coates RJ, Freedman DS, Williamson DF, Byers T. Do obese  
children become obese adults? A review of the literature. *Prev Med*, 22: 167-77, 1993.
- 3 Dietz WH. Health consequences of obesity in youth: childhood predictors of adult  
10 disease. *Pediatrics*, 101: 518-25, 1998.
- 4 Parsons TJ, Power C, Logan S, Summerbell CD. Childhood predictors of adult  
obesity: a systematic review. *Int J Obes*, 23: S1-107, 1999.
- 15 5 Rolland-Cachera MF, Deheeger M, Bellisle F, Sempe M, Guilloud-Bataille M, Patois  
E. Adiposity rebound in children: a simple indicator for predicting obesity. *Am J Clin Nutr*,  
39: 129-35, 1984.
- 6 Taylor RW, Grant AM, Goulding A, Williams, SM. Early adiposity rebound: review  
20 of papers linking this to subsequent obesity in children and adults. *Curr Opin Clin Nutr Metab  
Care*, 8: 607-12, 2005.
- 7 Baird J, Fisher D, Lucas P, Kleijnen J, Roberts H, Law C. Being big or growing fast:  
systematic review of size and growth in infancy and later obesity. *Br Med J*, 331: 929, 2005.
- 25 8 Monteiro PO, Victora CG. Rapid growth in infancy and childhood and obesity in later  
life--a systematic review. *Obes Rev*, 6: 143-54, 2005.
- 9 Ong KK, Loos RJ. Rapid infancy weight gain and subsequent obesity: systematic  
30 reviews and hopeful suggestions. *Acta Paediatr*, 95: 904-8, 2006.
- 10 Stettler N. Nature and strength of epidemiological evidence for origins of childhood  
and adulthood obesity in the first year of life. *Int J Obes*, 31: 1035-43, 2007.
- 35 11 Uppsala University. Uppsala family study. <http://www.pubcare.uu.se/family>, viewed  
16 October 2007.
- 12 Leon DA, Koupil I, Mann V, Tuvemo T, Lindmark G, Mohsen R, et al. Fetal,  
developmental, and parental influences on childhood systolic blood pressure in 600 sib pairs:  
40 the Uppsala Family Study. *Circulation*, 112: 3478-85, 2005.
- 13 Karlberg J, Luo ZC, Albertsson-Wikland K. Body mass index reference values (mean  
and SD) for Swedish children. *Acta Paediatr*, 90: 1427-34, 2001.
- 45 14 Ruppert D, Wand MP, Carroll RJ. *Semiparametric Regression*. Cambridge University  
Press, New York, 2003.
- 15 Goldstein H. *Multilevel Statistical Models*. Hodder Arnold, London, 2003.

- 16 Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*, 38: 963-74, 1982.
- 17 Durbán M, Harezlak J, Wand MP, Carroll RJ. Simple fitting of subject-specific curves  
5 for longitudinal data. *Stat Med*, 24: 1153-67, 2005.
- 18 Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child*, 73: 25-9, 1995.
- 10 19 Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB, the Avon Longitudinal Study of Pregnancy and Childhood Study Team. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *Br Med J*, 320: 967-71, 2000.
- 15 20 Fuentes RM., Notkola I-L, Shemeikka S, Tuomilehto J, Nissinen A. Tracking of body mass index during childhood: a 15-year prospective population-based family study in eastern Finland. *Int J Obes*, 27: 716-21, 2003.
- 21 Charney E, Goodman HC, McBride M, Lyon B, Pratt R. Childhood antecedents of adult obesity. Do chubby infants become obese adults? *N Engl J Med*, 295: 6-9, 1976.  
20
- 22 Robinson GK. That BLUP is a good thing: the estimation of random effects. *Statistical Science*, 6: 15-51, 1991.

## Tables

**Table 1 Distributions of variables at birth and at physical examination in later childhood for the 1162 subjects with at least 3 body mass index (BMI) observations between birth and age 3 years, by sex.**

Variable	Males ( <i>n</i> = 596)			Females ( <i>n</i> = 566)		
	Mean	Median	SD	Mean	Median	SD
<b>At birth</b>						
Weight (kg)	3.74	3.73	0.60	3.65	3.71	0.56
Length (cm)	51.6	51.0	2.2	50.9	51.0	2.2
<b>At physical examination in later childhood (age 5-13 years)</b>						
Age (years)	10.0	10.1	1.7	10.1	10.2	1.8
BMI z-score	0.26	0.12	1.20	0.36	0.30	1.09

**Table 2 Distributions of age and body mass index (BMI) at BMI peak, by sex and birthweight group. The percentage of subjects with identified BMI peak was calculated as a percentage of those included in each subgroup model.**

Sex	Birthweight group <sup>1</sup>	No. (%) with BMI peak	Age at BMI peak (years)			BMI at BMI peak (kg/m <sup>2</sup> )		
			Mean	Median	SD	Mean	Median	SD
Males	Concordant high	102 (83.6%)	0.79	0.78	0.13	18.5	18.3	1.4
	Discordant	289 (86.3%)	0.72	0.67	0.17	18.1	18.1	1.3
	Concordant low	126 (90.6%)	0.72	0.65	0.16	17.7	17.6	1.3
	Total	517 (86.7%)	0.73	0.69	0.16	18.1	18.0	1.4
Females	Concordant high	121 (83.4%)	0.76	0.70	0.20	17.9	17.9	1.3
	Discordant	272 (90.7%)	0.79	0.75	0.17	17.9	17.7	1.2
	Concordant low	118 (97.5%)	0.87	0.88	0.13	17.1	17.0	1.2
	Total	511 (90.3%)	0.80	0.76	0.17	17.7	17.7	1.3

5 <sup>1</sup> Siblings had ‘concordant high’ or ‘concordant low’ birthweight if both were in the top or bottom quarter of the birthweight distribution, respectively. Siblings had ‘discordant birthweight’ if the sex-adjusted difference in birthweight between them was 0.4 kg or more.

**Table 3 Estimated effects, 95% confidence intervals (CI) and Wald test P-values for the models for body mass index (BMI) z-score at physical examination fitted on age or BMI at BMI peak, birthweight and sex. Models were additionally adjusted for age at physical examination. Models were fitted on the 1028 subjects with an identified BMI peak.**

Explanatory variable	Coefficient	95% CI	P-value
Age at BMI peak (years)	0.94	0.54, 1.34	<0.001
Birthweight (kg)	0.47	0.35, 0.60	<0.001
Female sex	0.10	-0.02, 0.23	0.1
BMI at BMI peak (kg/m <sup>2</sup> )	0.33	0.28, 0.39	<0.001
Birthweight (kg)	0.18	0.06, 0.31	0.003
Female sex	0.27	0.15, 0.39	<0.001

5

**Table 4 Estimated effects, 95% confidence intervals (CI) and Wald test P-values for the models for body mass index (BMI) z-score at physical examination fitted jointly on age and BMI at BMI peak, birthweight and sex. Model was additionally adjusted for age at physical examination. Model was fitted on the 1028 subjects with a successfully identified BMI peak.**

Explanatory variable	Coefficient	95% CI	P-value
Age at BMI peak (years)	0.59	0.21, 0.97	0.003
BMI at BMI peak (kg/m <sup>2</sup> )	0.32	0.27, 0.37	<0.001
Birthweight (kg)	0.23	0.10, 0.35	<0.001
Female sex	0.23	0.11, 0.35	<0.001

5

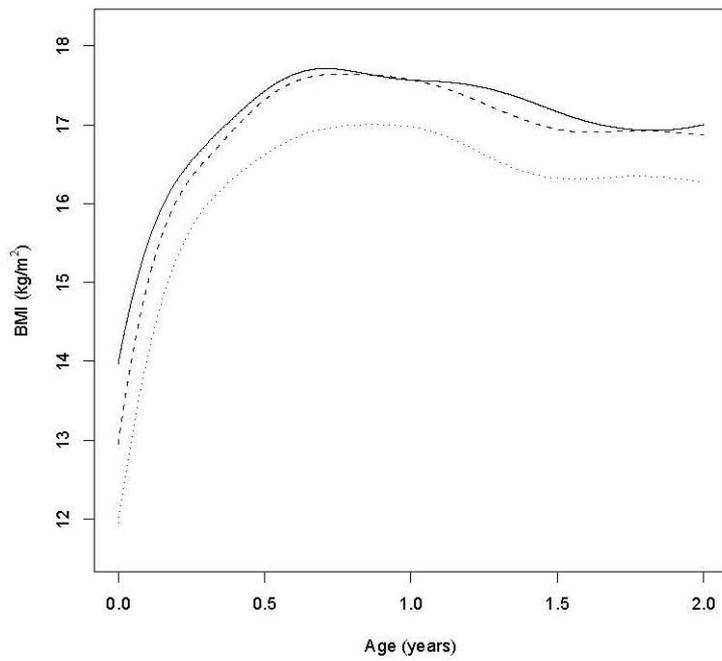
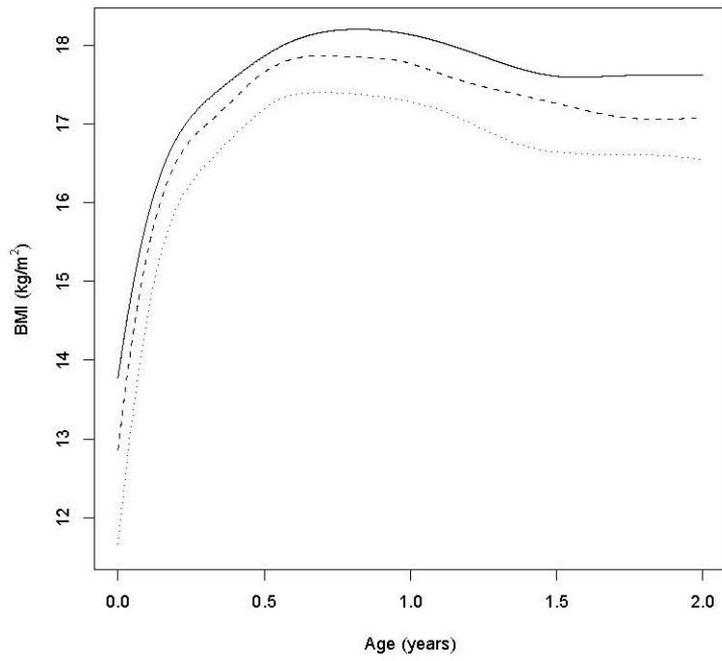


Fig. 1. Population average curves for males (upper plot) and females (lower plot). Solid line is the concordant high birthweight group, dashed line is the discordant birthweight group, dotted line is the concordant low birthweight group.

5

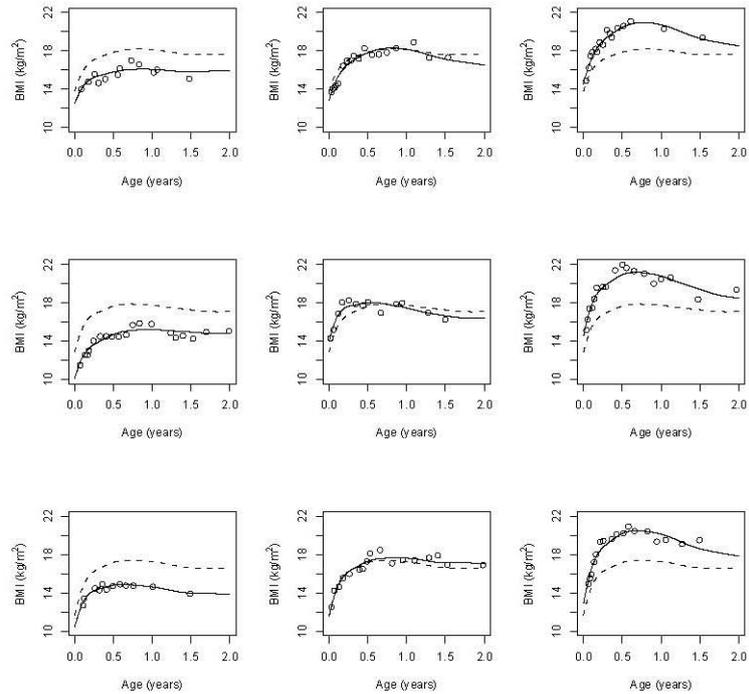


Fig. 2. Observed BMI values (circles), estimated population average curves (dashed lines) and fitted subject-specific curves (solid lines) for nine males. The top row corresponds to the concordant high birthweight model, the middle row to the discordant birthweight model and the bottom row to the concordant low birthweight model. Within each row the left hand plot is for a subject who had generally low body mass index (BMI) through infancy, the middle plot is for a subject who had average BMI and the right hand plot is for a subject who had high BMI. Within each combination of birthweight group and BMI level the plotted subject is chosen at random from those eligible.

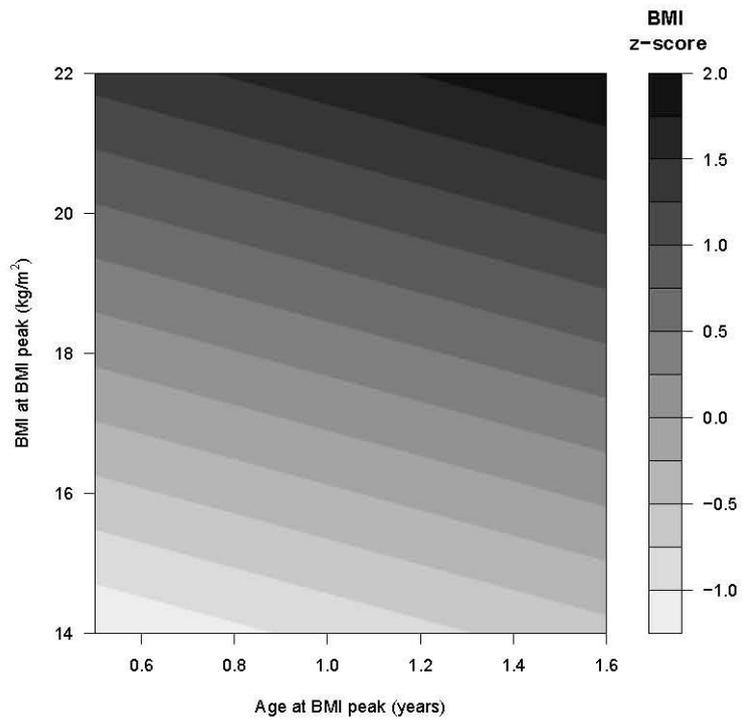


Fig. 3. Contour plot for predicted body mass index (BMI) z-score in later childhood for different combinations of age and BMI at BMI peak in males with mean birthweight and mean age at physical examination.