Joint Multivariate Response Modeling for Repeated BMI Measures and Single Measures of Adult Cardiovascular Disease Risk Factors

Leah Li
MRC Centre of Epidemiology for Child Health, University College London, 30 Guilford Street, London WC1N 1EH, United Kingdom

Abstract
Studies of the association between developmental trajectories and adult health require methods relating developmental measures from different life stages to single measures of health outcomes. We present a joint multivariate response model to a longitudinal response variable, body mass index (BMI), and two single measure adult outcomes - systolic blood pressure (SBP) and high density lipoprotein cholesterol **(**HDL-C), to investigate the association between BMI trajectories and adult cardiovascular disease (CVD) risk factors. We adopt a linear spline model for repeated BMI measures to allow for distinct childhood and adult curves and separate models for SBP and HDL-C. The models are fitted simultaneously by assuming the joint distribution of random coefficients. The model is applied to the 1958 British Birth Cohort (n=17,000), whose BMI was recorded at six ages from 7 to 45y (16,820 with one or more measures) and SBP and HDL-C were measured at 45y. Results show that the rate of BMI gain in adulthood has a stronger association with SBP and HDL-C than the rate of BMI growth in childhood (p<0.05 for SBP in boys and for HDL-C in both genders). For SBP, the estimated correlation for the rate of adult BMI gain is 0.27 (95% CI: 0.23, 0.32) in males and 0.35 (0.31, 0.38) in females, compared to 0.22 (0.16, 0.27) and 0.18 (0.11, 0.25) respectively for the rate of childhood growth. For HDL-C the correlation is -0.43 (-0.48, -0.38) and -0.45 (-0.48, -0.41) for adult BMI gain compared to -0.14 (-0.17, -0.10) and -0.26 (-0.30, -0.22) for childhood growth. Furthermore, the rate of childhood growth is associated with adult outcomes, independent of adult BMI gain.
Conclusions Joint multivariate response modeling is a useful approach for estimating the association between repeated exposure variables at different life stages and adult outcomes, therefore has important applications in the life-course epidemiology.

Key Words: child-to-adult BMI trajectories, CVD risk factors, joint multivariate modeling, life-course, repeated measures, multilevel model

1. Introduction

In longitudinal birth cohort studies, variables are often collected for each individual at multiple time points. Cohort studies are often used in life course epidemiology to investigate how risks for adult diseases may accumulate over time and to identify critical periods when the effect has a greater impact. For example, physical growth and development in childhood and adulthood have been associated with adult disease risks. A simple approach used to study such an association is a conditional linear regression model where all repeated growth measures are included as independent variables. The model is re-parameterized in terms of the first measure and all subsequent consecutive increments [6, 14, 17]. The conditional models have been applied to assess how changes of BMI across the life course influence later cardiovascular disease (CVD) risk factors [4, 17, 20]. This approach does not account for the fact that timing of measures may not be the same across individuals and it includes only individuals with complete data on the repeated measurements, although multiple imputation may be applied for the missing growth measures [6]. The conditional model may become impractical when there are a large number of growth measurements as there would be many consecutive intervals. Some have suggested that we only include time points representing important periods of growth instead of using all available measures [4]. Another two-step approach is to first fit a random effect model to the repeated growth measurements. The individual-specific growth parameters (i.e. slope, which indicates the rate of growth) are estimated from the model and are then used as independent variables in the primary model for each disease risk [1, 5, 18]. As the individual-specific parameters are estimated, they are subject to errors. This approach requires the assumption that enough information is available for each individual to obtain the estimates of the growth parameters and their variances with adequate precision.

We consider a flexible approach which is closely related to the general multivariate normal linear random effects model [15]. We use a joint model for a mixture of single and repeated measurements of response variables to examine the relationship between longitudinal changes of BMI and adult CVD risk factors. We adopt a linear spline model with random coefficients for BMI to allow for different curves for BMI changes in childhood and adulthood, and separate models for CVD risk factors with only a random intercept term. The models for BMI and CVD risk factors are fitted simultaneously by considering the joint distribution of their random coefficients. The joint model can be defined in a multilevel framework, where the measurement are assumed to be level-1 units, clustered within individuals (level-2 units) [11]. The estimates of growth characteristics are derived from the model's parameters. The correlations between growth characteristics and values of CVD risk factors are obtained directly using the estimated covariance matrix. Thus these joint models have important applications to life-course study of adult health, for example, to explore how parameters characterize growth and development at different life stages are associated with adult disease risks.

The outline of this paper is as follows: Section 2 describes the use of a joint model to analyze longitudinal BMI measurements and two adult CVD risk factors: adult systolic blood pressure (SBP) and high density lipoprotein cholesterol (HDL-C); it also describes the correlation coefficients between growth characteristics of BMI (i.e. BMI at specific ages or rates of BMI changes) and each CVD risk factor. Section 3 describes the 1958 British birth cohort study and the results from the joint model analysis, and Section 4 contains a discussion.

2. A joint model for single and longitudinal response variables

2.1 The joint model

In the present case of one longitudinal (BMI) and two single measure response variables (SBP and HDL-C), we denote $y_{ij}^{(1)}$ to be the log 10 transformed BMI ($y_{ij}^{(1)} = \log BMI_{ij}$, $j = 1,2,..n$, $i = 1,2,...,n_j$, n_j is the number of measurements on individual j), t_{ij} to be the exact age at each measurement, $y_j^{(2)}$ and $y_j^{(3)}$ to be adult SBP and log 10 transformed HDL-C ($y_j^{(3)} = \log HDL_j$). We assume that $y_{ij}^{(1)}$ follows a piecewise linear curve with a single knot at age t_0 (same for all individuals). The indicator represents two age ranges, with $I_{t_{ij}>t_0} = 1$ for $t_{ij} > t_0$ or $I_{t_{ij}>t_0} = 0$ otherwise. We specify only a random intercept term for $y_j^{(2)}$ and $y_j^{(3)}$. The joint model for multivariate response variables can be written as

$$\begin{aligned} y_{ij}^{(1)} &= \beta_{0j} + \beta_{1j} t_{ij} + \beta_{2j}(t_{ij} - t_0) I_{t_{ij}>t_0} + e_{ij} \\ y_j^{(2)} &= \beta_{3j} \\ y_j^{(3)} &= \beta_{4j} \end{aligned} \qquad (1)$$

where random coefficients β_{0j}, β_{3j} and β_{4j} are individual-specific intercepts for the three response variables. For the response variable $y_{ij}^{(1)}$, there are two individual-specific slopes on age, β_{1j} (slope 1 for $t_{ij} \le t_0$) and $\beta_{1j} + \beta_{2j}$ (slope 2 for $t_{ij} > t_0$), which indicate the degree of BMI changes over time. The coefficients $(\beta_{0j}, \beta_{1j}, ..., \beta_{4j})$ characterize individual's deviation from the estimated mean values. Model (1) is fitted jointly assuming that random coefficients $\boldsymbol{\beta_j} = (\beta_{0j}, \beta_{1j}, \beta_{2j}, \beta_{3j}, \beta_{4j})^T$ are independently and identically distributed (iid) across *j* and follow a multivariate normal distribution with mean $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)^T$ and individual-level (level-2) covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} & \sigma_{03} & \sigma_{04} \\ & \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ & & & \sigma_3^2 & \sigma_{34} \\ & & & & \sigma_4^2 \end{bmatrix}.$$

An unstructured covariance matrix is specified. The errors e_{ij} are assumed to be iid across *i* and *j* with $e_{ij} \sim N(0, \sigma_e^2)$, and independent of the random effects $\boldsymbol{\beta_j}$. In the simultaneous equations model (1), the responses are dependent on each other. The covariance terms represent the degree to which the individual intercept and slopes for the response variable $y_{ij}^{(1)}$ (β_{0j}, β_{1j}, β_{2j}) co-vary within themselves and with values of $y_j^{(2)}$ and $y_j^{(3)}$ (β_{3j} and β_{4j}).

Model (1) can be expressed as a multivariate model $\boldsymbol{y_j} = \boldsymbol{X_j^T \beta_j} + \boldsymbol{\varepsilon_j}$, where $\boldsymbol{y}_j^T = (y_j^{(1)}, y_j^{(2)}, y_j^{(3)})^T$ is an $(n_j + 2) \times 1$ vector of response variables on individual *j* and $\boldsymbol{y}_j^{(1)} = (y_{1j}^{(1)}, y_{2j}^{(1)}, ... y_{n_j j}^{(1)})^T$, β_j is a 5×1 vector of random effects, $\boldsymbol{X}_j^T$ is an $(n_j + 2) \times 5$ matrix

$$X_j^T = \begin{bmatrix} Z_j & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

where Z_j is an $n_j \times 3$ matrix of constants

$$Z_j = \begin{bmatrix} 1 & t_{1j} & (t_{1j} - t_0) \cdot I_{t_{1j} > t_0} \\ 1 & t_{2j} & (t_{2j} - t_0) \cdot I_{t_{2j} > t_0} \\ . & . & . \\ 1 & t_{n_j j} & (t_{n_j j} - t_0) \cdot I_{t_{n_j j} > t_0} \end{bmatrix}$$

and $\boldsymbol{\varepsilon_j} = (e_{1j}, e_{2j}, ..., e_{n_j j}, 0, 0)^T$ is an $(n_j + 2) \times 1$ vector of level-1 errors.

2.2 Correlations between growth characteristics of $y_{ij}^{(1)}$ and values of $y_j^{(2)}$ and $y_j^{(3)}$

The growth characteristics we are interested include the values of $y_{ij}^{(1)}$ at age *t*, written as $\beta_{0j} + \beta_{1j} t + e_{ij}$ $(t \le t_0)$ and $\beta_{0j} - \beta_{2j} t_0 + (\beta_{1j} + \beta_{2j}) t + e_{ij}$ $(t > t_0)$, and slopes for $y_{ij}^{(1)}$, slope 1 (β_{1j}) and slope 2 ($\beta_{1j} + \beta_{2j}$). The correlation between values of $y_{ij}^{(1)}$ at age *t* and $y_j^{(2)}$ (i.e. β_{3j}), which is a function of t, can be written as

$$R(t) = \frac{\sigma_{03} + \sigma_{13} t + \sigma_{23} (t - t_0) \cdot I_{t>t0}}{\sqrt{[\sigma_0^2 + 2\sigma_{01} t + \sigma_1^2 t^2 + \sigma_2^2 (t - t_0)^2 \cdot I_{t>t0} + 2(\sigma_{02} + \sigma_{12} t)(t - t_0) \cdot I_{t>t0} + \sigma_e^2] \cdot \sigma_3^2}} \quad (2)$$

The correlation between each slope of $y_{ij}^{(1)}$ and the values of $y_j^{(2)}$ can be expressed as follows

$$\frac{\sigma_{13} + \sigma_{23} \cdot I_{t>t0}}{\sqrt{[\sigma_1^2 + (2\sigma_{12} + \sigma_2^2) \cdot I_{t>t0}] \cdot \sigma_3^2}} \quad (3)$$

The correlation coefficient (between -1 and 1) assumes that the measures (slope of $y_{ij}^{(1)}$ and intercept $y_j^{(2)}$**)** follow a normal distribution.

There is no constraint added to the covariance matrix, thus formula (3) represents the correlation between each slope of $y_{ij}^{(1)}$ and $y_j^{(2)}$ values when the correlation between the other slope and $y_j^{(2)}$ values is accounted for through the covariance matrix. For example, $\sigma_{13} / \sqrt{\sigma_1^2 \cdot \sigma_3^2}$ is the correlation between slope 1 of $y_{ij}^{(1)}$ $(t \le t_0)$ and $y_j^{(2)}$ values when the correlation between slope 2 $(t > t_0)$ and $y_j^{(2)}$ values is adjusted for. We may also estimate the correlation between slope 1 of $y_{ij}^{(1)}$ and $y_j^{(2)}$ when the correlation for slope 2 of $y_{ij}^{(1)}$ is not accounted for. This can be estimated by adding a constraint of $\sigma_{13} + \sigma_{23} = 0$ to the covariance matrix. Similarly, we can estimate the correlation between slope 2 and $y_j^{(2)}$ values ignoring the correlation between slope 1 and $y_j^{(2)}$ by setting a constraint $\sigma_{13} = 0$. This application is particularly useful in life-course epidemiology to study whether BMI increase during a specific period affects an outcome, independent of BMI changes during other periods.

For the response variable $y_j^{(3)}$, its correlations with growth characteristics of $y_{ij}^{(1)}$ can be obtained by replacing the parameters for $y_j^{(2)}$ in (2) and (3) with the respective parameters for $y_j^{(3)}$.

We fit the joint model in MLwiN. The model's parameters (β , Σ , and σ_e^2) are first estimated using the iterative generalized least squares (IGLS) algorithm of Goldstein (1986) which provides maximum likelihood estimates [21]. For the linear spline model for $y_{ij}^{(1)}$, the knot is positioned at different ages and -2log (likelihood) of each model is compared. The location of the knot t_0 is chosen based on the likelihood profile function [13].

The correlation coefficients are non-linear functions of the parameter estimates. Thus their theoretical distributions are complicated and the explicit formulae of their variances can not be directly expressed. There are several approaches to calculate the variance for the estimate that is a non-linear function of model parameters, for example using Taylor series [7] and re-sampling methods [19]. In the current study we apply the semi-parametric bootstrap re-sampling procedure. Each sample is drawn (with replacement) from estimated residuals at each level to create a dataset that has the same multilevel structure. These sampled residuals are then added to the fixed part of the model to obtain a new set of responses. Estimates of model parameters and the correlation coefficients are obtained for each of the bootstrap sample. We re-sampled 499 bootstrap datasets to obtain the means and 95% confidence intervals (CI) for the correlations coefficients.

3. Child-to-adult BMI and adult CVD risk factors in the 1958 British birth cohort

3.1 Data

The 1958 British birth cohort includes all children born in England, Wales, and Scotland in one week, March 1958 [2]. Approximately 17,000 live births were followed-up at multiple ages from 7 to 50y. Height and weight were recorded in childhood at 7, 11, and 16y, and in adulthood at 23, 33, and 45y. All variables were measured by medical personnel using standard protocols, except for 23y when height and weight were self-reported. Body mass index (BMI, kg/m^2) was derived at each age. At 45y, 12,069 cohort members were invited to participate in a medical assessment. Blood pressure (BP, mmHg) was measured three times using an Omron 705 automated sphygmomanometer. The average of three systolic blood pressure (SBP) measurements is used in the analyses. Non-fasting venous blood samples were collected. Total and high density lipoprotein cholesterol (HDL-C, mmol/L) were analyzed using enzymatic methods. The formation on anti-hypertensive medication was recorded. The cohort remained in adulthood has been found to be representative of the original birth sample in most respects [9]. Ethical approval for data collection at 45y was given by the South East England Multi-Centre Research Ethics Committee.

The main analysis involves 8,657 male and 8,163 female cohort members with at least one BMI measure between 7 and 45y, or a measurement of SBP or HDL-C at 45y. Thus a majority of participants are included in the analysis (all individuals had at least one BMI measure, 9297 had a SBP measure and 7808 had a HDL-C measure). We apply model (1) to three response variables, longitudinal BMI measurements from 7 to 45y, and individual level SBP and HDL-C. The aim is to examine how BMI at particular ages and changes of BMI from childhood to adulthood are associated with SBP and HDL-C at 45y. As in model (1), BMI at each age and HDL-C are log 10 transformed to correct for their right skewness. At age 45y, 429 individuals were taking anti-hypertensive medication. The estimated association between BMI trajectories and SBP is likely to be underestimated if the treatment effect is not accounted for. In the current analysis, we add a constant of 10mmHg to the observed SBP in treated participants [22]. BMI trajectories and measures of CVD risk factors differ by genders. The observed mean BMI and SBP at 45y are higher for males than females; 27.8 kg/m^2

(sd=4.3) versus 27.0 kg/m^2 (sd=5.5) for BMI, and 133.5 mmHg (sd=15.4) versus 121 mmHg (sd=16.0) for SBP. Mean HDL-C is lower for males than females; 1.43 mmol/l (sd=0.34) versus 1.69 mmol/l (sd=0.41).

There are a maximum of six BMI measures per person. The observed mean BMI shows distinct rates of BMI changes in childhood and adulthood (Figure 1).

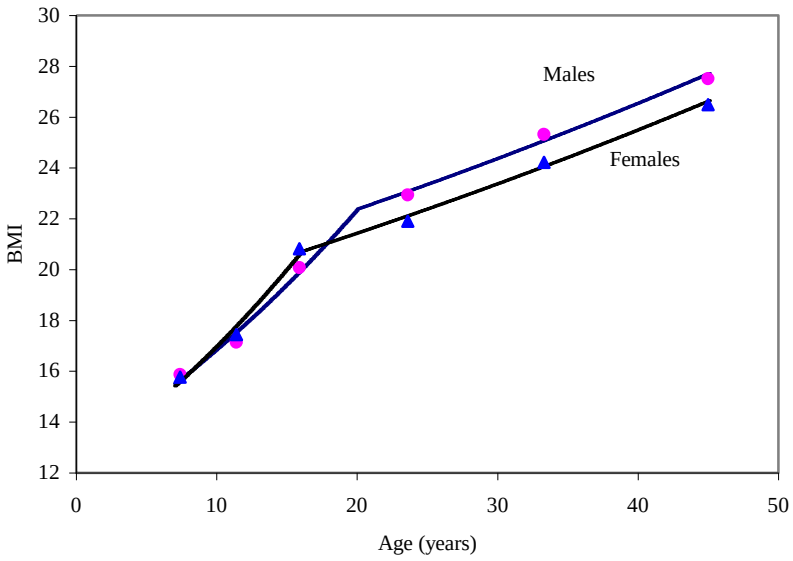

Figure 1: Geometric mean BMI trajectories (estimated from joint model) and observed values

We fit two linear curves to repeated BMI measurements, one for "childhood'' and one for "adulthood'', where the knot t_0 is the age at which the time rate of growth changes. A gender-specific knot is chosen at age 20y for males and 16y for females based on likelihood profile (Figures 2a and 2b). We fitted model (1) separately to males and females. The parameter estimates for fixed effects are given in Table 1. The correlation coefficients between growth characteristics and SBP and HDL-C at 45y are given in Table 2.

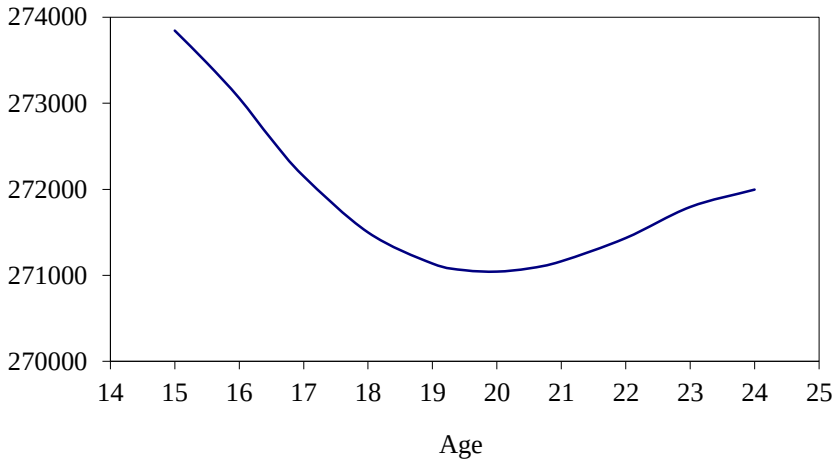

Figure 2a: -2loglikelihood (1958 males)

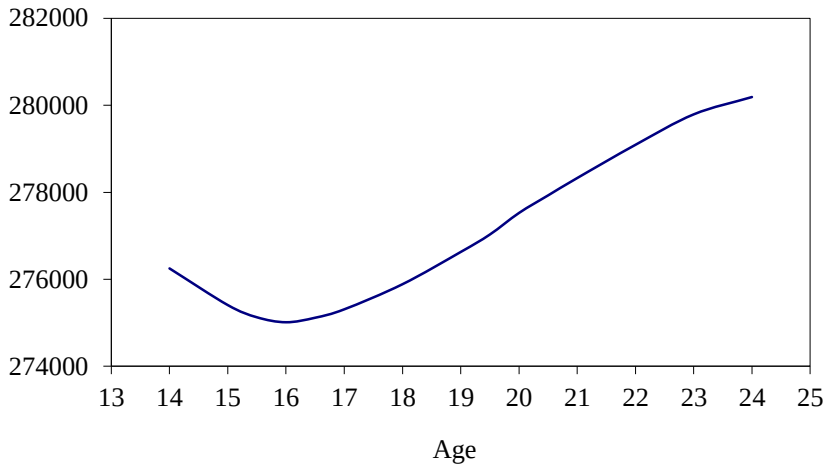

Figure 2b: -2loglikelihood (1958 females)

3.2 Results

Child-to-adult BMI

The curves of geometric mean BMI shown in Figure 1 are derived from the fixed part of the model for BMI ($10^{\beta_0+\beta_1 t+\beta_2(t-t_0)I_{t>t_0}}$). The estimated curves are close to the observed values. There is a significant difference between child and adult slopes, indicating that BMI increase is faster in childhood than in adulthood (Table 1, *p*-value for β_2 <0.05). Comparing the curves for males and females, there is little difference in mean BMI at 7y, but BMI trajectories diverge thereafter when girls gain BMI at a faster rate than boys in childhood. Childhood slope for $\log BMI$ is 0.0142 (se=0.0001) for girls, compared to 0.0121 (se=0.0001) for boys (*p*-value for the gender difference <0.05). Their

rate is similar in adulthood. The gender difference remains and geometric mean BMI for males (27.7 kg/m^2) is significantly higher than that for females (26.6 kg/m^2) (p-value for the difference <0.05).

Table 1 Parameter estimates (standard error) for BMI trajectories ($\log BMI$), SBP and HDL-C ($\log HDL$) at 45y using bootstrap sampling †

BMI trajectories	**Males**	**Females**
Childhood slope (β_1)	0.0121(0.0120,0.0122)	0.0142(0.0140, 0.0143)
Adulthood slope ($\beta_1+\beta_2$)	0.0038(0.0037, 0.0038)	0.0038(0.0037, 0.0039)
difference in slopes (β_2)	-0.0083(-0.0085,-0.0081)	-0.0104(-0.0106,-0.0102)
Geometric mean at 7	15.51(15.48,15.55)	15.42(15.37,15.46)
Geometric mean at 11	17.35(17.31,17.38)	17.56(17.52,17.61)
Geometric mean at 45	27.69(27.62,27.83)	26.63(26.47,26.77)
SBP		
Intercept for SBP (β_3)	133.48(133.03,133.93)	120.95(120.47,121.48)
HDL-C		
Intercept for $\log HDL$ (β_4)	0.1435(0.1401,0.1469)	0.2143(0.2109,0.2177)
Geometric mean HDL at 45	1.3916(-0.6162,3.3994)	1.638(-0.3628,3.6388)

† 499 bootstrap samples were draw to estimate the parameters and their standard errors

Child-to-adult BMI trajectories and SBP and HDL- cholesterol at 45 years

The mean SBP, estimated from the fixed part of the model, is close to the observed values. The geometric mean for HDL-C is estimated at 1.39 mmol/L for males and 1.64 mmol/L for females (Table 1). There is only a weak inverse correlation between SBP and HDL-C (Table 2).

Table 2 Correlation coefficients (95% CI) - Estimated using bootstrap sampling †

	Bootstrap estimate (95% CI)	
SBP	**Males**	**Females**
BMI7 & SBP	0.03(-0.00, 0.05)	0.03(-0.05,0.05)
BMI11 & SBP	0.07(0.04, 0.10)	0.06(0.04,0.08)
BMI45 & SBP	0.27(0.24, 0.29)	0.29(0.27,0.32)
Child slope & SBP	0.22(0.16, 0.27)	0.18(0.11, 0.25)
Adult slope & SBP	0.27(0.23, 0.32)	0.35(0.31, 0.38)
HDL-cholesterol		
BMI7 & HDL-C	-0.044 (-0.056, -0.031)	-0.097 (-0.110,-0.081)
BMI11 & HDL-C	-0.068 (-0.087, -0.048)	-0.148 (-0.168, -0.124)
BMI45 & HDL-C	-0.287 (-0.316, -0.258)	-0.414 (-0.448, -0.376)
Child slope & HDL-C	-0.135 (-0.174, -0.097)	-0.263 (-0.299, -0.219)
Adult slope & HDL-C	-0.429 (-0.482, -0.380)	-0.447 (-0.479, -0.411)

† 499 bootstrap samples were draw to estimate the parameters and their standard errors

Using the covariance estimates, the correlation coefficients between BMI from age 7 to 45y and SBP and HDL-C, a function of age, are estimated from expression (2). The correlations are also given by Figure 3 for SBP and Figure 4 for HDL-C to illustrate how the correlations change over age. SBP at 45y is weakly associated with BMI at 7y, but the correlation strengthens with increasing age (Figure 3). Thus SBP has a strong correlation with concurrent BMI at 45y, 0.27 (95% CI: 0.24, 0.29) for males and 0.29 (95% CI: 0.27, 0.32) for females (Table 2). The findings indicate that higher adult SBP is associated with larger increases in BMI from 7y onwards.

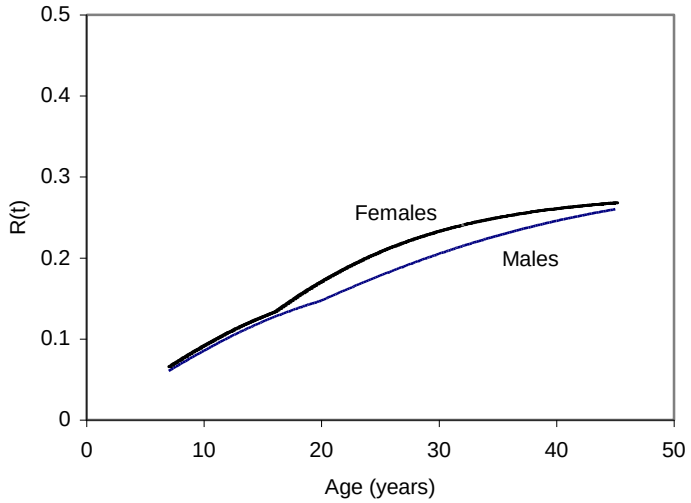

Figure 3: Correlation coefficients between log *BMI* and SBP

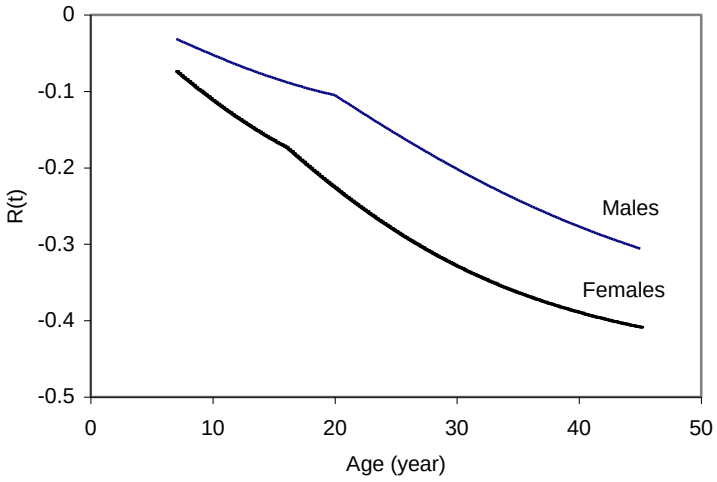

Figure 4: Correlation coefficients between log *BMI* and log *HDL*

Applying expression (3) we obtain the estimates for correlation coefficients between the rates of BMI changes and SBP. Large BMI increases in childhood and adulthood are both significantly correlated

with adult SBP. When the correlations of child and adult slopes with SBP at 45y are both considered in the model, the correlation for child slope with SBP is 0.22 (0.16, 0.27) for males and 0.18 (0.11. 0.25) for females, and the correlation for adult slope with SBP is higher, at 0.27 (0.23, 0.32) and 0.35 (0.31, 0.38) respectively (Table 2). When only one slope is considered (i.e. the correlation between the other slope and SBP is set to be zero), the correlation between child slope and SBP becomes stronger, at 0.25 (0.23, 0.28) and 0.20 (0.16, 0.24) respectively (data not presented). The correlation between adult slope and SBP is 0.35 (0.31, 0.39) and 0.36 (0.33, 0.40). This indicates that although SBP has a stronger association with rate of BMI gain in adulthood, rate of BMI growth in childhood is significantly associated with adult SBP, independent of rate of BMI gain in adulthood.

Figure 4 shows the correlation between $\log BMI$ from 7 to 45y and $\log HDL$ at 45y. Patterns of relationships for HDL-C are similar to those for SBP, but in the opposite direction. Although BMI at 7y has a weak inverse correlation with HDL-C, the correlation also strengthens with increasing age (Figure 4). The correlation with HDL-C is much stronger with rate of BMI change in adulthood than in childhood. Similar to SBP, the negative correlations between BMI changes in childhood and adulthood and HDL-C at 45y remain significant when correlation between child and adult slopes and HDL-C are both considered (Table 2). This indicates that rates of BMI increases in childhood and adulthood are independently associated with HDL-C at 45y.

4. Discussion

We fitted joint models to repeated BMI measurements and SBP and HDL-C in mid-adulthood to investigate the associations between life-course BMI trajectories and adult CVD risk factors in the 1958 cohort. Joint models can be applied to a wide variety of situations, including the life-course study of adult diseases. The key feature of joint modeling is that we have an individual-level covariance matrix for the longitudinal and single measure response variables. We can obtain the estimates for the correlation coefficients between growth characteristics of the longitudinal variable and later outcomes using the components of the covariance matrix.

The application of this class of joint models to life-course studies has several advantages: first, individuals do not have to be measured at the same ages or have the same number of time points, thus subjects with incomplete data are not excluded from the analysis; second, the model takes into account of the fact that measurements are clustered within individuals (i.e. between repeated measures and between multiple response variables); and third, the model is useful when there is a large number of growth measurements and simple approaches examining consecutive intervals are impractical.

Goldstein has shown a joint multivariate response model for predicting adult height using repeat childhood height and bone age [11] and more recently predicting adult BMI and glucose level study from growth between birth and 10 years of age [12]. The multivariate response model has also been used to study the effect of school resources on pupil attainment [21]. Our aim is not to make prediction of adult outcomes, but to use the estimation of the variance-covariance components of the random coefficients to quantify the strength of the associations between characteristics of child-to-adult BMI trajectories and SBP and HDL-C at 45y. Williams (2001) has used a joint multivariate model proposed by Zucker [24] to longitudinal BMI and blood pressure measurements in children [23]. In the present study, we adopt a linear spline model with random coefficients to repeated BMI measurements to allow for distinct growth parameters for childhood and adulthood. A two-step approach has been used in previous research where the individual-specific growth parameters estimated from a growth curve model in the first stage are used as predictors of adult outcomes in the second stage [1, 3, 8, 10, 18].

The second-stage model however ignores the fact that the growth parameters are sample estimates and therefore subject to sampling variability. One important feature of the joint model presented here is that individual-specific growth characteristics are directly related to each disease risk to estimate their associations. The model is flexible as it allows to jointly model not only responses of the same type (like our example of BMI and CVD risk factors), but also responses of different types [12]. However in the joint model we do not assume a causal relationship between BMI trajectories and adult outcome.

In the 1958 cohort, there are relatively small number (a maximum of six) of BMI measures on individuals between 7 and 45y. The measures are widely spaced over the life-course and at each follow up there is little spread. Thus we assume a linear spline model with one knot, which describes the average rate of growth over two time periods (childhood and adulthood) [16], but can not capture the detailed patterns of BMI changes (i.e. critical period such as pubertal growth spurt). When there are more measurements are available, the joint model can be extended to include more knots or more complex function of age to characterize growth trajectories. In the current study, we have also explored similar joint models but applying quadratic spline curves to repeated BMI measures. Although these models provide a smooth curve with a continuous growth rate, the data do not support the quadratic spline model as the current model provides the better fit for the data (Figures 5a and 5b). Covariates can be added to the joint models for child-to-adult growth or adult outcomes. The correlation coefficients would be interpreted as the association between BMI growth and SBP (or HDL-C) that is not explained by the covariates.

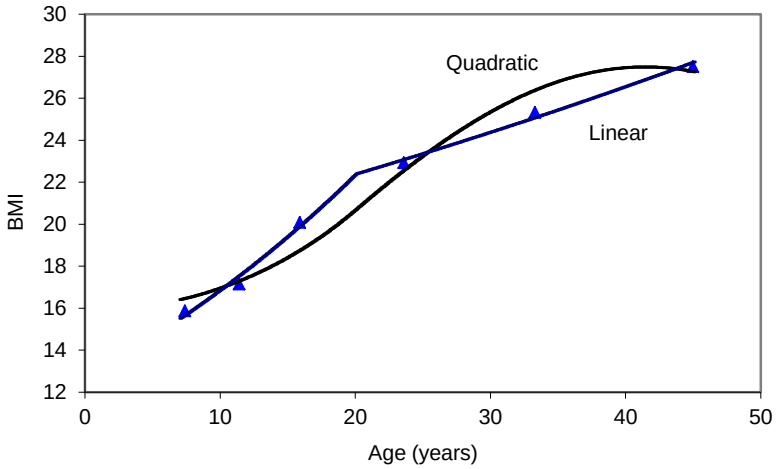

Figure 5a: Geometric mean BMI trajectories (linear and quadratic spline curves) and observed mean values for males

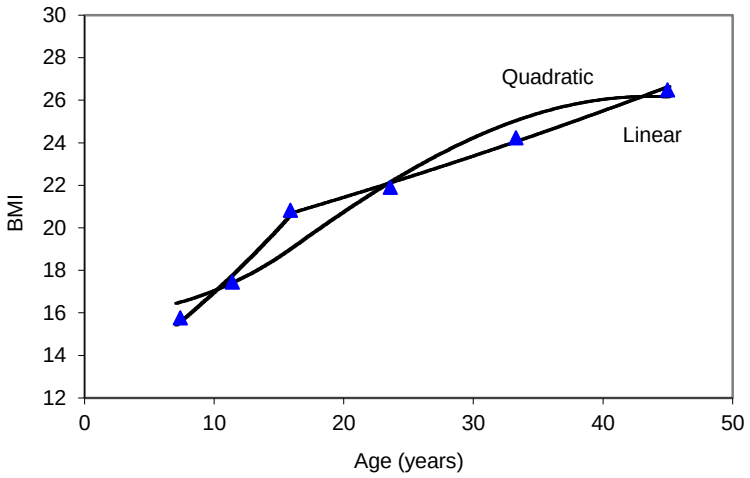

Figure 5b: Geometric mean BMI trajectories (linear and quadratic spline curves) and observed mean values for females

There are critical periods, for example, during infancy and puberty when growth may be important in relation to CVD risks in late life. High adult BMI has been associated with adverse CVD risk factors such as raised BP and lower HDL-C [20]. Few studies have investigated the long term effect of childhood growth on CVD risks in later life. Using a joint model for repeated BMI measures and SBP and HDL-C at 45y, we found that allowing for the within individual correlation, although BMI at 7y has only a weak association with SBP and HDL-C, larger increases in BMI subsequently both during child and adult life, which accumulate to a higher adult BMI, are strongly associated with a higher SBP and a lower HDL-C in mid-adulthood. Although rate of BMI change is greater in childhood than in adulthood, BMI gain in adulthood has a stronger association with CVD risk factors than BMI growth in childhood. This is expected as BMI increase in childhood is part of normal development, whereas BMI increase in adulthood is due almost entirely to weight gain, as adult height changes little once attained. Although there is only a weak correlation between SBP and HDL-C, high BMI and rate of BMI increase both in childhood and adulthood are significantly associated with higher SBP and lower HDL-C. Prevalence of obesity continues to increase in many populations with younger generations becoming obese at an earlier age. Our findings suggest that control of excessive weight gain at any life stage will have a beneficial effect on reducing adult CVD risks.

Acknowledgements

Leah Li is funded as an MRC Career Development Fellow in Biostatistics (grant code: G0601941). Research at the Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust benefits from R&D funding received from the NHS Executive. The author would like to thank Professor Harvey Goldstein for useful comments on the draft.

Reference List

1. Ben-Shlomo Y, McCarthy A, Hughes R, Tilling K, Davies D, Smith GD. Immediate postnatal growth is associated with blood pressure in young adulthood: the Barry Caerphilly Growth Study. *Hypertension* 2008; **52**(4):638-644.
2. Centre for Longitudinal Studies Institute of Education. National Child Development Study Composite File including selected Perinatal Data and sweeps one to five [computer file]. National Birthday Trust Fund, National Children's Bureau, City University Social Statistics Research Unit [original data producers]. Colchester,Essex: The Data Archive [distributor]. SN:3148. 1994.
3. Cheung YB, Low L, Osmond C, Barker D, Karlberg J. Fetal growth and early postnatal growth are related to blood pressure in adults. *Hypertension* 2000; **36**(5):795-800.
4. Cole TJ. Modeling postnatal exposures and their interactions with birth size. *J Nutr* 2004; **134**(1):201-204.
5. De Stavola BL, Hardy R, Kuh D, Silva IS, Wadsworth M, Swerdlow AJ. Birthweight, childhood growth and risk of breast cancer in a British cohort. *Br J Cancer* 2000; **83**(7):964-968.
6. De Stavola BL, Nitsch D, dos SS, I, McCormack V, Hardy R, Mann V, et al. Statistical issues in life course epidemiology. *Am J Epidemiol* 2006; **163**(1):84-96.
7. Demnati A, Rao JNK. Linearization variance estimators for survey data. *Survey Methodology* 2004; **30**:17-26.
8. dos Santos Silva I, De Stavola BL, Mann V, Kuh D, Hardy R, Wadsworth ME. Prenatal factors, childhood growth trajectories and age at menarche. *Int J Epidemiol* 2002; **31**(2):405-412.
9. Ferri E. *Life at 33:the fifth follow-up of the National Child Development Study*. London: National Children's Bureau; 1993.
10. Fraser A, Hughes R, McCarthy A, Tilling K, Davies D, Rumley A, et al. Early life growth and hemostatic factors: the Barry Caerphilly Growth study. *Am J Epidemiol* 2008; **168**(2):179-187.
11. Goldstein H. Efficient statistical modeling of longitudinal data. *Ann Hum Biol* 1986; **13**:129-141.
12. Goldstein H, Kounali D. Multilevel multivariate modelling of childhood growth, number of growth measurements and adult characteristics. *J R Statist Soc A* 2009; **172**:599-613.
13. Hall CB, Ying J, Kuo L, Lipton RB. Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Computational Statistics & Data Analysis* 2003; **42**:91-109.
14. Hardy R, Wadsworth ME, Langenberg C, Kuh D. Birthweight, childhood growth, and blood pressure at 43 years in a British birth cohort. *Int J Epidemiol* 2004; **33**(1):121-129.
15. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; **38**(4):963-974.
16. Li L, Hardy R, Kuh D, Lo CR, Power C. Child-to-adult body mass index and height trajectories: a comparison of 2 British birth cohorts. *Am J Epidemiol* 2008; **168**(9):1008-1015.
17. Li L, Law C, Power C. Body mass index throughout the life-course and blood pressure in mid-adult life: a birth cohort study. *J Hypertens* 2007; **25**(6):1215-1223.
18. McCarthy A, Hughes R, Tilling K, Davies D, Smith GD, Ben-Shlomo Y. Birth weight; postnatal, infant, and childhood growth; and obesity in young adulthood: evidence from the Barry Caerphilly Growth Study. *Am J Clin Nutr* 2007; **86**(4):907-913.
19. Shao J, Tu D. *The jackknife and bootstrap*. New York: Springer-Verlag; 1995.
20. Skidmore PM, Hardy RJ, Kuh DJ, Langenberg C, Wadsworth ME. Life course body size and lipid levels at 53 years in a British birth cohort. *J Epidemiol Community Health* 2007; **61**(3):215-220.
21. Steele F, Vignoles AJA. The effect of school resources on pupil attainment: a multilevel simultaneous equation modeling. *J R Statist Soc A* 2007; **170**:801-824.
22. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* 2005; **24**(19):2911-2935.

23. Williams S. Modelling the association between increases in height and adiposity and changes in blood pressure between ages 7 and 18 years. Klein B, Korsholm L. Proceeding of the 14th International Workshop on Statistical Modelling , 521-524. 2001. Odense, Denmark.
24. Zucker DM, Zerbe GO, Wu C. Inference for the association between coefficients in a multivriate growth model. *Biometrics* 1995; **51**:413-423.