

RESEARCH ARTICLE

Open Access



# Using linear and natural cubic splines, SITAR, and latent trajectory models to characterise nonlinear longitudinal growth trajectories in cohort studies

Ahmed Elhakeem<sup>1,2\*</sup> , Rachael A. Hughes<sup>1,2</sup>, Kate Tilling<sup>1,2</sup>, Diana L. Cousminer<sup>3,4,5</sup>, Stefan A. Jackowski<sup>6,7</sup>, Tim J. Cole<sup>8</sup>, Alex S. F. Kwong<sup>1,2,9</sup>, Zheyuan Li<sup>10,11</sup>, Struan F. A. Grant<sup>3,4,5,12,13</sup>, Adam D. G. Baxter-Jones<sup>6</sup>, Babette S. Zemel<sup>12,14</sup> and Deborah A. Lawlor<sup>1,2</sup>

## Abstract

**Background:** Longitudinal data analysis can improve our understanding of the influences on health trajectories across the life-course. There are a variety of statistical models which can be used, and their fitting and interpretation can be complex, particularly where there is a nonlinear trajectory. Our aim was to provide an accessible guide along with applied examples to using four sophisticated modelling procedures for describing nonlinear growth trajectories.

**Methods:** This expository paper provides an illustrative guide to summarising nonlinear growth trajectories for repeatedly measured continuous outcomes using (i) linear spline and (ii) natural cubic spline linear mixed-effects (LME) models, (iii) Super Imposition by Translation and Rotation (SITAR) nonlinear mixed effects models, and (iv) latent trajectory models. The underlying model for each approach, their similarities and differences, and their advantages and disadvantages are described. Their application and correct interpretation of their results is illustrated by analysing repeated bone mass measures to characterise bone growth patterns and their sex differences in three cohort studies from the UK, USA, and Canada comprising 8500 individuals and 37,000 measurements from ages 5–40 years. Recommendations for choosing a modelling approach are provided along with a discussion and signposting on further modelling extensions for analysing trajectory exposures and outcomes, and multiple cohorts.

**Results:** Linear and natural cubic spline LME models and SITAR provided similar summary of the mean bone growth trajectory and growth velocity, and the sex differences in growth patterns. Growth velocity (in grams/year) peaked during adolescence, and peaked earlier in females than males e.g., mean age at peak bone mineral content accrual from multicohort SITAR models was 12.2 years in females and 13.9 years in males. Latent trajectory models (with trajectory shapes estimated using a natural cubic spline) identified up to four subgroups of individuals with distinct trajectories throughout adolescence.

**Conclusions:** LME models with linear and natural cubic splines, SITAR, and latent trajectory models are useful for describing nonlinear growth trajectories, and these methods can be adapted for other complex traits. Choice of

\*Correspondence: [a.elhakeem@bristol.ac.uk](mailto:a.elhakeem@bristol.ac.uk)

<sup>1</sup> MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

Full list of author information is available at the end of the article



method depends on the research aims, complexity of the trajectory, and available data. Scripts and synthetic datasets are provided for readers to replicate trajectory modelling and visualisation using the R statistical computing software.

**Keywords:** ALSPAC, Bone mineral content, BMDCS, Growth models, Life-course, Mixed-effects, PBMAS, Tutorial

## Background

Appropriate modelling of repeated measures in cohort studies can improve our understanding of: (i) patterns of change across the life-course (e.g., developmental trajectories to peak function and age-related decline); (ii) influences on these patterns of change; and (iii) influence of variation in patterns of change on later health and well-being [1]. Many developmental processes display non-linear patterns of change with age, especially during the growing years, which makes it important but challenging to accurately model their trajectories [2]. Also requiring attention is the choice of method to appropriately address the research question, e.g., whether to use methods that model an average trajectory in the whole sample [3–5] or clustering based approaches to identify groups of individuals with similar trajectories [6]. Moreover, there is a lack of accessible and practical guidance which can discourage novice researchers from using more complex methods [7]. Thus, an overview of sophisticated modelling procedures along with open-source software applications and application in multiple different cohorts can help address these challenges.

This paper provides a guide to describing nonlinear longitudinal growth trajectories for a single repeatedly-measured continuous outcome using linear and natural cubic regression splines [3, 4], SITAR (Super Imposition by Translation and Rotation) models [5], and (spline-based) latent trajectory models [6] – all common methods for examining growth. The next section of the paper gives an overview of modelling nonlinear growth and the various models considered. The four approaches (and the appropriate interpretation of their results) are then illustrated by modelling bone mass trajectories across three cohort studies to characterise patterns of change and their sex differences. The final section provides recommendations about when different modelling methods might be useful and discusses approaches and challenges in analysing exposures and outcomes of patterns of change, and in making cross-cohort comparisons.

## Methods

### Modelling nonlinear growth

A variety of statistical methods are available for handling repeated (correlated) observations from the same individuals and analysing trajectories [8]. Most methods involve fitting growth models within either a structural

equation modelling framework (e.g., latent growth curve analysis where change is analysed as a latent process [9]) or a multilevel modelling framework, with both giving similar results under certain conditions [10]. One type of repeated measures model that is useful when the primary interest is estimating a population-average trajectory is generalised estimating equations (GEE) [11, 12]. This uses a working covariance matrix to correct for the dependence among repeated observations but is usually not suited for examining variation within/between individuals. Another repeated measures model that can estimate both population-average and individual-specific trajectories (and is more robust to missing outcome data than GEE) is the mixed-effects model [13].

### Mixed-effects models

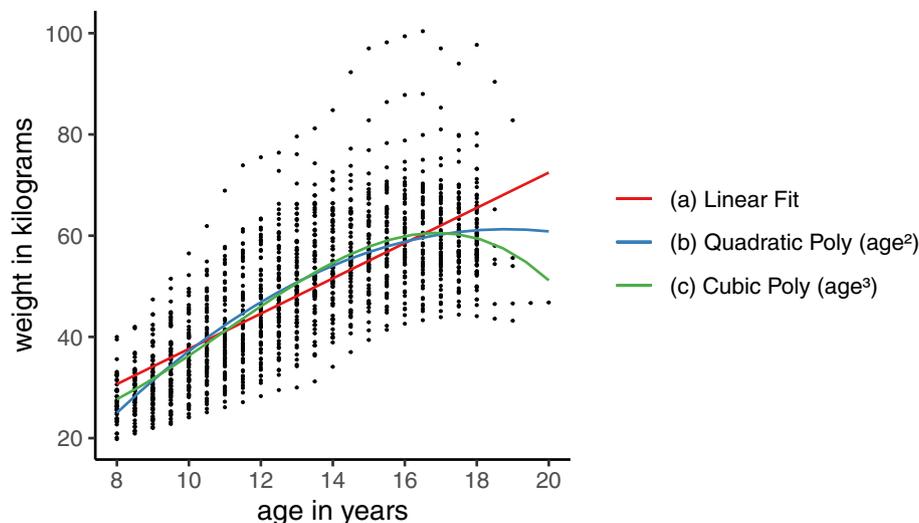
Mixed-effects models (random-effects, multilevel, or hierarchical models) estimate a population-average trajectory as ‘fixed effects’ and variation of individual trajectories around this average as ‘random effects’ [12–14]. A common form is the linear mixed-effects (LME) model where the repeated outcome is modelled by a linear combination of the fixed and random effects. An LME model for a single continuous outcome (e.g., weight), as a *linear* function of time, which includes random intercepts and random slopes can be written as follows:

$$y_{ij} = \beta_{0i} + \beta_{1i}t_{ij} + \varepsilon_{ij}, \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2), i.i.d. \quad (1)$$

$$\beta_{0i} = \beta_0 + u_{0i}, u_{0i} \sim N(0, \sigma_0^2), i.i.d. \quad (1.1)$$

$$\beta_{1i} = \beta_1 + u_{1i}, u_{1i} \sim N(0, \sigma_1^2), i.i.d. \quad (1.2)$$

where,  $y_{ij}$  denotes a single outcome  $y$  measured in individual  $i$  ( $i=1, 2, \dots, N$ ) at time  $t_{ij}$  ( $j=1, 2, \dots, J_i$ ), with responses ( $y_1 \dots y_N$ ) assumed to be independent between individuals.  $\beta_{0i}$  and  $\beta_{1i}$  are individual-specific intercept and slope terms (respectively) that have fixed effects ( $\beta_0, \beta_1$ ) and random effects ( $u_{0i}, u_{1i}$ ). The random effects  $u_i$  are assumed to be independently normally distributed with mean zero and covariance matrix  $\Omega_u$ . Residual errors  $\varepsilon_{ij}$  are assumed to be independently identically normally distributed with variance  $\sigma_\varepsilon^2$  and reflect the difference between observed and predicted values for



**Fig. 1** Example illustrating the limitations of using polynomial functions to approximate a nonlinear growth trajectory. Coloured lines represent predicted trajectories from LME models with age as (a) linear term and as (b) quadratic polynomial and (c) cubic polynomial. Points display weight measurements taken from 70 females in the Berkeley Child Guidance Study. Dataset was originally provided as an appendix to the book by Tuddenham and Snyder (1954). The data for this example were taken from the freely accessible 'Berkeley' dataset provided with the 'sitar' package [15]

individual  $i$  at occasion  $j$ . The random effects  $u_i$  and residuals  $\varepsilon_{ij}$  are assumed to be mutually independent.

### Moving beyond a linear trajectory

The LME model in Eq. (1) assumes linear change in the outcome with increasing time (e.g., age). Nonlinear change, which is common (particularly when modelling change over a large age range and/or during periods of rapid complex growth) can be incorporated into LME models by including linear combinations of nonlinear terms for age in the model – i.e., keeping the linear link function. Historically, the standard approach has been to make use of polynomial functions to approximate nonlinear curves. However, as illustrated in Fig. 1, polynomials have limitations, e.g., simpler polynomials give few curve shapes and more complex polynomials tend to fit badly at extremes and produce artefactual turns in the curve. A more flexible alternative to analysing complex patterns of change (e.g., with several peaks/troughs, as with body mass index (BMI) over infancy, childhood, and adolescence) is using a set of connected polynomials covering different segments of the time/age distribution known as spline functions.

### Spline functions

A spline function is a set of piecewise polynomials that are joined together at break/turning points called knots (see [16] for an introductory review). Splines are formed using spline basis functions (the dimensional space

containing each element from a set of local polynomials) and can be fitted by various methods including linear and nonlinear models. The polynomial degree and method of knot placement is generally what distinguishes between different spline functions. There are three main ways of estimating spline functions: as regression splines, penalised regression splines, and smoothing splines [12, 16–19]. Regression splines use fewer knots than observations – here both the number and location of knots must be specified by the user. Penalised regression splines are regression splines that include an added penalty on the parameters – here, the user does not need to specify the knots, instead, a (maximum) spline basis complexity (flexibility) is specified, and the spline coefficients are then penalised by a tuning parameter to avoid overfitting. Smoothing splines use a knot at every unique observation with penalisation of the estimated curve – these preceded penalised regression splines and are now less widely used.

In this paper, we consider two commonly used types of regression splines: linear splines and natural cubic splines. These can be parameterised by a linear combination of the transformed covariate (e.g., age) and thus they can be fitted within an LME framework to model a repeated outcome as a nonlinear function of time. Fitting penalised regression spline models is covered briefly in the Discussion section. The next three sections give an overview to using linear and natural cubic splines in LME models and approaches to choosing the number

and position of knots. SITAR, which utilises a natural cubic spline within a nonlinear mixed effects framework, is then covered. These three mixed-effects modelling approaches all estimate population-average trajectories (and individual variation from these) and so can be used when it is hypothesised that the study population shares a common mean trajectory. Extensions to latent trajectory models (with trajectory shapes estimated using natural cubic splines) for situations where the aim/hypothesis is to identify suspected latent heterogenous trajectories in subpopulations as opposed to one mean trajectory for the whole population are covered in the subsequent section. All four modelling approaches are summarised in Table 1.

**Linear spline LME models**

The simplest spline function is the linear spline which can be used to describe growth by a series of connected lines joined at knots, where the slope can change after each knot [3]. For example, a linear spline for age (measured from 5 to 40 years) with 2 knots at ages 10 and 15 years produces three different linear slopes of the repeated outcome measure (e.g., weight): 5 to ≤10, 10 to ≤15, and 15 to ≤40 years. The LME model in Eq. (1) can be rewritten to include a linear spline function for age  $b(t)$  with  $K$  knots  $\xi_1 < \xi_2 < \dots < \xi_K$  as:

$$y_{ij} = \beta_{0i} + \beta_{1i}t_{ij} + \sum_{k=1}^K b_{ki}(t_{ij} - \xi_k)_+ + \varepsilon_{ij} \quad (2)$$

$$(t_{ij} - \xi_k)_+ = \begin{cases} 0 & t_{ij} < \xi_k \\ (t_{ij} - \xi_k) & t_{ij} \geq \xi_k \end{cases} \quad (2.1)$$

The model in Eq. (2) includes a linear spline in both the fixed and random effects (with  $b_{ki}$  having fixed effect  $b_k$  and random effect  $v_{ki}$ ) which allows for nonlinear mean and individual trajectories, respectively. The fixed and random splines are assumed to have the same knots in Eq. (2) however, it is possible to allow fewer knots in the random spline. If the aim was solely to model a nonlinear mean trajectory, then Eq. (2) can be simplified by replacing the random spline with a random line – here,  $\beta_{0i}$  and  $\beta_{1i}$  have similar interpretation as in Eq. (1), with random effect  $v_{ki}$  omitted (i.e.,  $b_{ki} = b_k$ ). The rate of change of a linear spline (1st derivative) is not continuous at the knots. An alternative function that has continuous 1<sup>st</sup> and 2<sup>nd</sup> derivatives is the natural cubic spline.

**Natural cubic spline LME models**

A natural cubic spline (also known as restricted cubic spline) is a set of cubic polynomials with continuity and slope constraints at each knot, and additional constraint of linearity at the extremes of the curve, typically before

the first and after the last knot [12, 16, 20, 21]. This extra linearity constraint makes the trajectory less erratic at the ends of the age distribution and so more reliable than linear splines (and unrestricted cubic splines). The inclusion of cubic terms with continuity constraints means natural cubic spline models can be more parsimonious for complex shapes than a linear spline with many knots. An LME model that includes a natural cubic spline function  $b(t)$  with  $K$  knots  $\xi_1 < \xi_2 < \dots < \xi_K$  (and a linearity constraint for values  $t < \xi_1$  and  $t > \xi_K$ ) can be written as:

$$y_{ij} = \beta_{0i} + \beta_{1i}t_{ij} + \sum_{k=1}^{K-2} b_{ki}(t_{ij} - \xi_k)_*^3 + \varepsilon_{ij} \quad (3)$$

$$(t_{ij} - \xi_k)_*^3 = (t_{ij} - \xi_k)_+^3 - (t_{ij} - \xi_{k-1})_+^3 \frac{\xi_k - \xi_{k-1}}{\xi_k - \xi_{k-1}} + (t_{ij} - \xi_k)_+^3 \frac{\xi_{k-1} - \xi_k}{\xi_k - \xi_{k-1}}, k = 1, 2, \dots, K - 2 \quad (3.1)$$

In words, a natural cubic spline for age (measured from 5 to 40 years) with 2 knots at 10 and 15 years (and with first and last knots at 5 and 40 years, respectively), invokes 3 cubic polynomials: between 5 to ≤10, 10 to ≤15, and 15 to ≤40 years, and has its curvature equal to 0 at ages 5 and 40 years. If the first and last knots were placed at older and younger ages respectively, then the curve would be linear from the first and last knot to the youngest and oldest ages, respectively. Note, we have defined the natural cubic spline using truncated power basis (like how the linear spline was defined). The analysis software creates this function using a B-spline basis which is mathematically challenging to represent but numerically more stable. Whatever basis is used, if the polynomial degree and knots are identical, then the spline will always be the same.

**Choosing number and location of knots**

The flexibility of regression splines is determined by the number/position of the knot points. For a natural cubic spline, the number of knots (rather than their position) is more important [20]. A small number of knots (between 3 and 5 knots) provides a good fit to some patterns [20] though, with many repeats (e.g., data spanning many decades) more knots may be required. Approaches to selecting the number/position of knots include (i) placing knots at quantiles of the age distribution (ii) using equally spaced knots, (iii) inspecting smoothing curves and using these to select knots, (iv) starting with many knots and reducing their number, and (v) placing knots at the mean age of data collection [3].

Model selection can be done informally (i.e., by inspecting plots of fitted values/residuals from competing models with different knots). Valid comparison between models with different knots (i.e., models with non-nested

**Table 1** Overview of linear spline LME models, natural cubic spline LME models, SITAR, and latent trajectory models for analysing nonlinear growth trajectories of a single repeatedly measured continuous outcome

	Linear spline LME model	Natural cubic spline LME model	SITAR	Latent trajectory model
Description	linear mixed-effects model with a linear spline function of the independent time variable	linear mixed-effects model with a restricted cubic spline function of the independent time variable	nonlinear mixed-effects model based on the shape invariant growth model	heterogenous growth curves fit to unknown subgroups of individuals
Advantages	easy to interpret the spline slope coefficients; can describe growth rate during different periods of the growth process	continuous 1 <sup>st</sup> & 2 <sup>nd</sup> derivatives give smoother trajectory and can identify points of peaks/troughs; linearity constraint gives a more reliable trajectory shape as less erratic at the tails of distribution	has useful features of the natural cubic spline, easy to estimate individual growth features – most notably individual ages at peak growth velocity	can identify unobserved sub-groups of individuals sharing distinct growth trajectories if any exist
Limitations	biologically implausible sudden changes in velocity (i.e., at the knots); erratic at the tails; cannot identify points of velocity maxima/minima; position (and location) of knots important	coefficients difficult to interpret (so plotting is more useful); can be challenging to estimate the individual growth curves due to complex spline basis functions used by the statistical software	may not work well for complex growth patterns e.g., with multiple peaks and troughs or where the growth curve does not plateau in adulthood	difficult to identify the optimal number of sub-groups; may identify implausible sub-groups; trajectories tend not to replicate in other cohorts
R package(s)	lme4; lspline	lme4; splines	sitar	lcm, splines

All models can include all individuals with at least one observed outcome measure, with valid estimates obtained under the assumption of outcome data missing at random (MAR) depending on observed values of the outcome and/or covariates. Note all models assume no autocorrelation (in our example, there are wide enough gaps between measures to assume that here)

mean structures) can be done using likelihood-based information criteria (e.g., Bayesian Information Criterion value (BIC)) provided maximum likelihood (ML) estimation is used [22]. Note however, ML-based variance estimates are biased downwards [12] and so one approach is to compare models fitted by ML and by restricted maximum likelihood (REML). Cross-validation is also useful for model selection [23]. If knot position was a primary interest (e.g., testing sensitive periods), then topic knowledge can inform placement of knots [24].

### SITAR models

SITAR is a shape invariant nonlinear mixed-effects model [5, 25]. Whereas LME spline models are linear models that allow terms to describe a non-linear trajectory with age, nonlinear mixed-effects models are fundamentally nonlinear in the coefficients [12]. SITAR assumes that a study population has a common characteristic curve (fitted as fixed effects), which through shifting and scaling (by a set of 3 random effects) can be transformed into any individual curve. Following the notation in Cole et al [5, 26], a SITAR model for outcome  $y$  can be written as:

$$y_{ij} = \alpha_0 + \alpha_i + h\left(\frac{t - \beta_0 - \beta_i}{\exp(-\gamma_0 - \gamma_i)}\right) + \varepsilon_{ij} \quad (4)$$

where  $y_{ij}$  is the outcome measurement for individual  $i$  at age  $j$ ;  $\alpha_0$ ,  $\beta_0$ ,  $\gamma_0$  are fixed effects;  $\alpha_i$ ,  $\beta_i$ , and  $\gamma_i$  are random effects for the  $i$ th individual;  $h(t)$  is a natural cubic spline curve; and  $\varepsilon_{ij}$  are independent normally distributed errors. The 3 random effects describe the size ( $\alpha_i$ ), timing ( $\beta_i$ ), and intensity ( $\gamma_i$ ) of individual growth relative to the mean growth curve.  $\alpha_i$  adjusts for the differences in  $y$  and geometrically reflects individual shifts up or down (translation) in the mean curve;  $\beta_i$  adjusts for differences in the timing of peak growth in  $y$  and geometrically reflects left to right shifts (translation) in the mean curve; and  $\gamma_i$  adjusts for the duration of the growth spurt and geometrically corresponds to shrinking or stretching of the age scale and rotating the curve.

Note, a key difference from other mixed effects models, with or without a natural cubic spline mean curve, is that SITAR models growth on both the  $x$ - and  $y$ -axes – this allows differences in developmental age to be modelled. Common practice in selecting the best fitting SITAR models is to compare models with varying number of knots placed at quantiles of the age distribution for the spline curve [5]. The internal SITAR model structure is also customisable [15].

### Latent trajectory models

The spline and SITAR models described above assume that the population is homogenous and described at the

population level by a mean trajectory, with variability of individuals about this mean. As an alternative, latent trajectory models assume that there is a heterogeneous population composed of unknown subgroups (latent classes) of individuals, each characterised by a unique mean trajectory profile [6, 27, 28]. Therefore, latent trajectory modelling addresses a different research hypothesis where heterogeneity in the population-average (mean) trajectory is of key interest. Latent trajectory models aim to minimise within group variance and maximise between group differences so that individuals are more similar within groups than between groups. Each individual has a probability of belonging to each latent class and is assigned to the class with the highest probability. Class membership is defined using a latent discrete random variable, with membership probability described by a multinomial logistic model.

Several latent trajectory modelling approaches are possible (see [6] for a recent overview). These include models that ignore the longitudinal structure (known as longitudinal latent class analysis) and models assuming no variability between individuals within subgroups (known as latent class growth analysis or group-based trajectory models). Another approach that is a direct extension of standard mixed-effects models is a growth mixture model, which involves fitting multiple growth curves to subgroups of individuals that share a common trajectory. Following the notation in [6], the LME model in Eq. (1) can be rewritten as a growth mixture model as:

$$y_{ij|c} = \beta_{0i}^c + \beta_{1i}^c t_{ij} + \varepsilon_{ij}^c \text{ for } c = 1, \dots, C, \quad (5)$$

where  $C$  indicates number of latent classes, with probabilities  $p_c$ ,  $c = 1, \dots, C$ , with  $0 \leq p_c \leq 1$  and  $\sum_{c=1}^C p_c = 1$ . All other terms are defined as before but specifically for each class  $c$ . Growth can be parameterised as nonlinear, e.g., using a natural cubic spline curve in each class (assuming same number/position of knots). Class-specific covariances for individual-level error terms can be included, and both fixed and random effects can be class specific.

Model estimation is conditional on a pre-specified number of classes, with the optimal number of classes identified through a combination of approaches. These include assessing interpretability and plausibility of classes e.g., inspecting if trajectories show biologically plausible patterns and examining characteristics (e.g., socioeconomic position) of the classes [29], information criteria, entropy (statistic for class separation), and numerically meaningful sub-groups (e.g.,  $\geq 5\%$  class size). Models with  $>1$  class are prone to local maxima solutions (i.e., convergence to the best solution in a neighbourhood of the parameter space, rather than the

global maximum (largest loglikelihood)). This can be avoided by using different starting values [30].

### Illustrative example

Next, using data from three cohort studies, we demonstrate the application of the four modelling approaches described above to characterise bone mineral content (BMC) growth trajectories and their sex differences, including in population subgroups with potentially distinct trajectories.

### Bone mass through the life course

Bone mass in early life is thought to be an important determinant of fracture and osteoporosis risk in later life [31] however, few studies have described its developmental trajectory. Furthermore, sex differences in osteoporotic fracture are assumed to be due to menopause but may also reflect early sexual dimorphism in bone development. This is a timely exploration, given the availability of studies with repeated measurements of BMC (a marker of bone strength).

### Studies and measurements

Three studies with repeated BMC measurements from childhood to adulthood were included (Table 2, Additional file 1): the Avon Longitudinal Study of Parents and Children (ALSPAC) [32, 33], Bone Mineral Density in Childhood Study (BMDCS) [34], and Pediatric Bone Mineral Accrual Study (PBMAS) [35]. Total-body (excluding-head) BMC was measured in grams using whole-body Dual-Energy X-ray Absorptiometry (DXA)

scans. Of note, the studies used DXA devices from different manufacturers (Lunar vs. Hologic) which scale differently and are not interchangeable but repeat scans within studies were acquired on the same device. Individuals from each study were included if they had  $\geq 1$  measure of BMC and no missing data on age at DXA scan (in years) or sex. Analyses were restricted to white ethnicity because 2 cohorts were ethnically homogeneous [33, 35]. The final analysis samples comprised 3888 males and 4007 females in ALSPAC, 465 males and 488 females in BMDCS, and 112 males and 127 females in PBMAS. All studies had ethics approval and obtained parental or participant informed consent.

### Statistical analysis

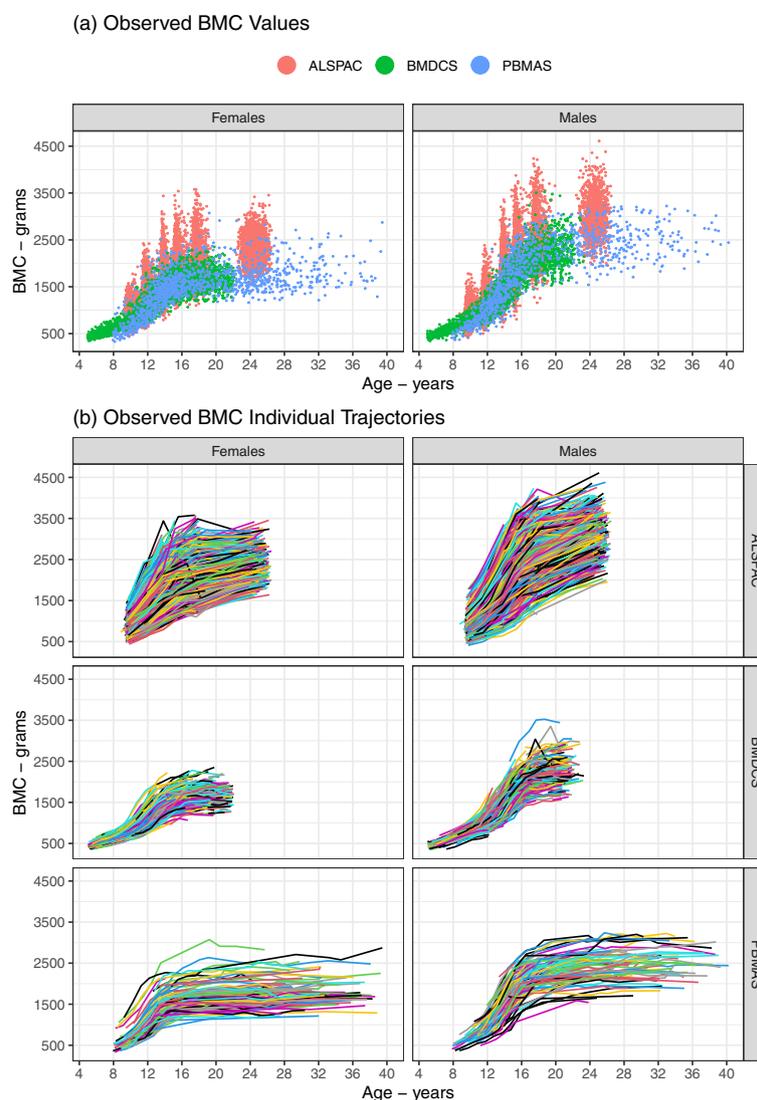
Analyses were performed in R version 4.0.2 (R Project for Statistical Computing) and RStudio version 1.3.1 (RStudio Team). R code is available at <https://github.com/aelhak/nltmr/>. Synthetic versions of the PBMAS cohorts were simulated [36] and can be found in the same repository.

Prior to trajectory modelling, scatterplots of BMC against age, and line plots of the individual BMC trajectories were used to inspect the form of the trajectory and identify clearly outlying observations. The datasets used for trajectory modelling (Fig. 2) showed as expected nonlinear change in BMC with age, and higher BMC in ALSPAC due to Lunar device. The numbers of individuals at each visit were described and age and BMC were summarised with means and standard deviations (Additional file 1). Models were fitted separately

**Table 2** Characteristics of the three cohort studies included in the trajectory modelling

Study name	Avon Longitudinal Study of Parents and Children (ALSPAC)	Bone Mineral Density in Childhood Study (BMDCS)	Pediatric Bone Mineral Accrual Study (PBMAS)
Design	birth cohort study (started in 1990–1992)	child cohort study (started in 2002–2003)	child cohort study (started in 1991)
Region and country	catchment area of 3 health authorities in Southwest England, UK	5 USA clinic centres: Los Angeles, New York, Cincinnati, Omaha, Philadelphia	2 elementary schools, Saskatoon, Saskatchewan, Canada
Birth years	1990–1992	1985–1997	1983–1976
Ethnicity	98% white ethnicity	ethnically diverse	95% white ethnicity
DXA device used to measure BMC	Lunar Prodigy	Hologic QDR-4500A	Hologic QDR-2000
Mean age at the baseline/youngest DXA scan (range)	9.9 years (8.8–11.7 years)	10.8 years (6.0–17.0 years)	11.8 years (8.0–15.1 years)
Mean age at last/oldest DXA scan (range)	24.6 years (22.4–26.5 years)	16.1 years (6.9–23.3 years)	37.3 years (34.3–40.2 years)
Frequency and the maximum number of repeated DXA scans	up to 6 repeated scans at mean ages 9.9, 11.7, 13.8, 15.4, 17.8, and 24.6 years	up to 7 yearly repeated scans	Up to 16 repeated scans (1991–1998, 2003–2005, 2007–2011 and 2016–2017)
Individuals included in the analysis <sup>a</sup>	4007 females 3888 males	488 females 465 males	127 females 112 males

<sup>a</sup>Trajectory modelling was restricted to white ethnicity individuals



**Fig. 2** Plots of the cohort datasets used in the trajectory modelling showing (a) bone mineral content (BMC) values at each age, and (b) BMC individual trajectories. Figure shows the observed BMC values by cohort and sex (a), and the observed BMC individual trajectories by cohort and sex (b)

by sex due to expected difference in bone mineral content accrual, and our aim was to explore this in the illustrative example.

Linear and natural cubic spline LME models were fitted using the ‘lme4’ package [37]. Models with 2 to 6 knots (placed at quantiles of the age distribution) in the fixed effects curve were compared. Nonlinear individual trajectories were allowed by including a random effects spline with 1 knot at the median. SITAR models with 2 to 6 knots (at quantiles of the age distribution) in the spline curve (and three random-effects) were fitted using the ‘sitar’ package [15]. LME and SITAR models were fitted by ML and best fitting models (optimum number

of knots) were determined by the smallest BIC (Additional file 2). Goodness of fit for the selected models was assessed by examining residuals (conditional on the random effects) from LME models and variance in BMC explained by SITAR models. The selected models were used to describe BMC growth trajectory and growth velocity. Slopes of the fixed effects spline segments from the linear spline models were used to summarise mean BMC velocity during different age windows and identify windows for peak growth. Mean peak BMC velocity and age at peak BMC velocity were obtained from the natural cubic spline LME and SITAR models by differentiating the mean spline curves.

Growth mixture models were fitted using the ‘*lcmm*’ package [30]. The forms of the best fitting mean natural cubic spline curves were used to model the fixed effects age curve. Models included random intercepts and random linear age slopes. Models with different numbers of latent classes were compared: from a 1-class model (i.e., a standard natural cubic spline LME model where all individuals follow a single mean trajectory) up to models with 5 classes. Models with > 1 class included class-specific random effects covariance matrices. An automatic search procedure was used to estimate each 2–5 class model for 100 iterations using random initial values from the distribution of the 1-class model. Optimal number of classes was chosen by inspecting predicted trajectory sub-groups from each model for biological plausibility, in addition to the smallest BIC and biggest entropy, and by excluding small class size ( $\geq 5\%$ ). Goodness of fit and discrimination capacity of the selected models (i.e. with optimal number of classes) was assessed by calculating posterior class membership probabilities [30].

## Results

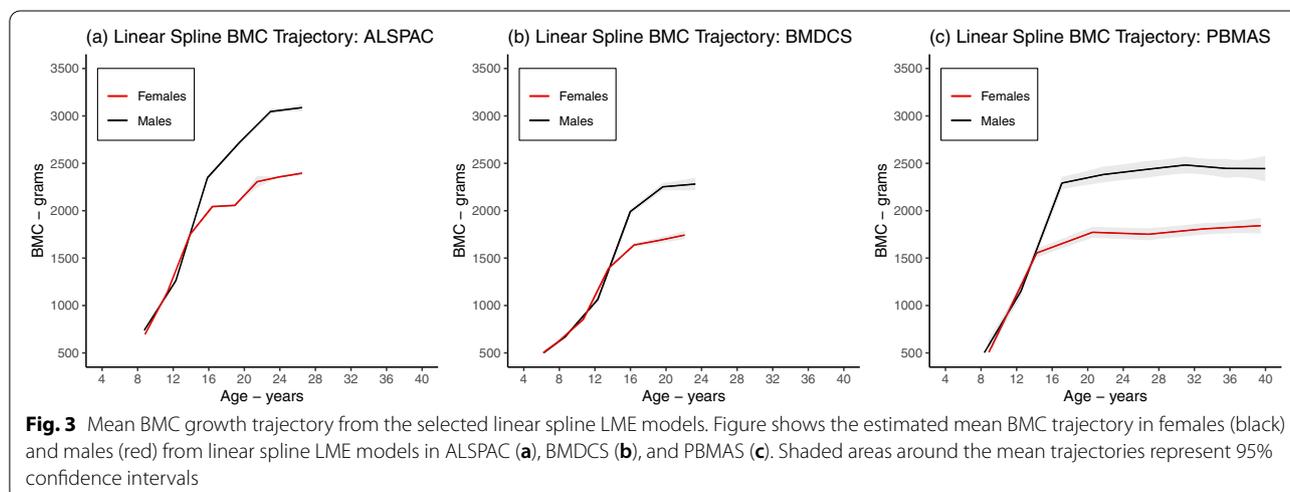
The mean predicted BMC trajectories in each cohort from the linear spline LME (Fig. 3), natural cubic spline LME (Fig. 4), and SITAR models (Fig. 5) showed that BMC increased with age up to a plateau in young adulthood, and thereafter remained stable to age 40 (in PBMAS). All models showed evidence of steeper growth trajectories through adolescence coinciding with emerging sex differences, with males subsequently having higher BMC than females, and plateauing later than females. BMC trajectories from all models were broadly similar (Fig. 6). Both linear and natural cubic spline LME models and SITAR provided a good fit to the data (Online Resource 2). The mean BMC growth velocity

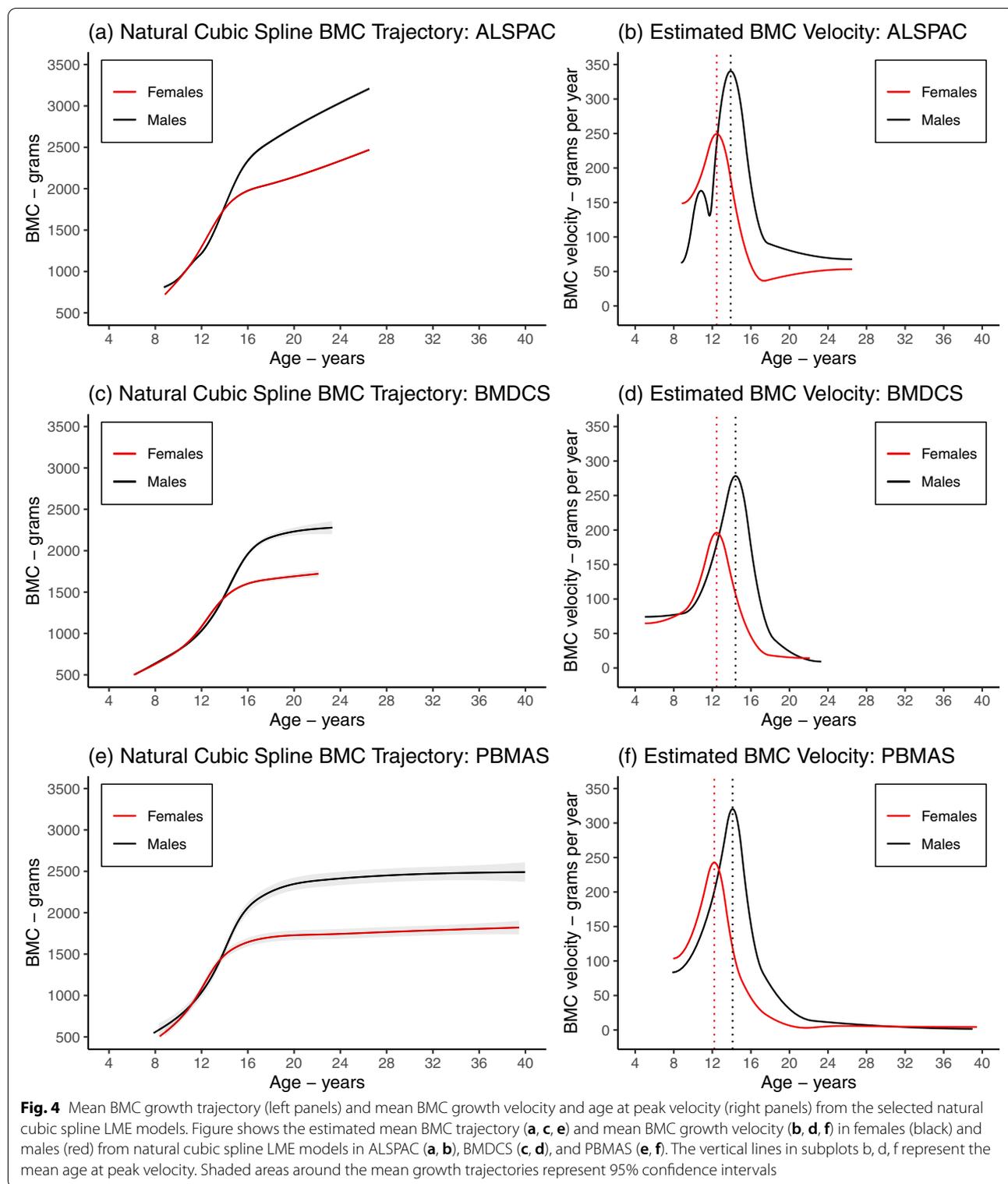
in different age windows (from linear spline LME models) peaked during adolescence, and the peak was lower and occurred earlier in females than males (Table 3). Mean BMC growth velocity curves from natural cubic spline LME and SITAR models were similar (Figs. 4–5): for all cohorts, the mean ages at peak BMC velocity from both models were within the age windows for peak growth identified by the linear spline LME models (Table 4).

The selected growth mixture models identified 3 subgroups in females and 2 subgroups in males from ALSPAC and BMDCS, and 4 in females and 3 in males from PBMAS (Fig. 7). Overall, differences in mean BMC between subgroups were larger during adolescence than in childhood and adulthood. One group with 15% of PBMAS females had higher mean BMC up to age 40 than the remaining three groups. A group comprising 27% of PBMAS males reached a lower peak and showed signs of bone loss by age 40, compared to the other two groups. The 2 trajectory groups in ALSPAC males displayed a biologically implausible lack of plateau by early adulthood. Model discrimination capacity was better in PBMAS (and BMDCS) than in ALSPAC, likewise entropy was high in PBMAS and low in ALSPAC and BMDCS (Additional file 3).

## Discussion

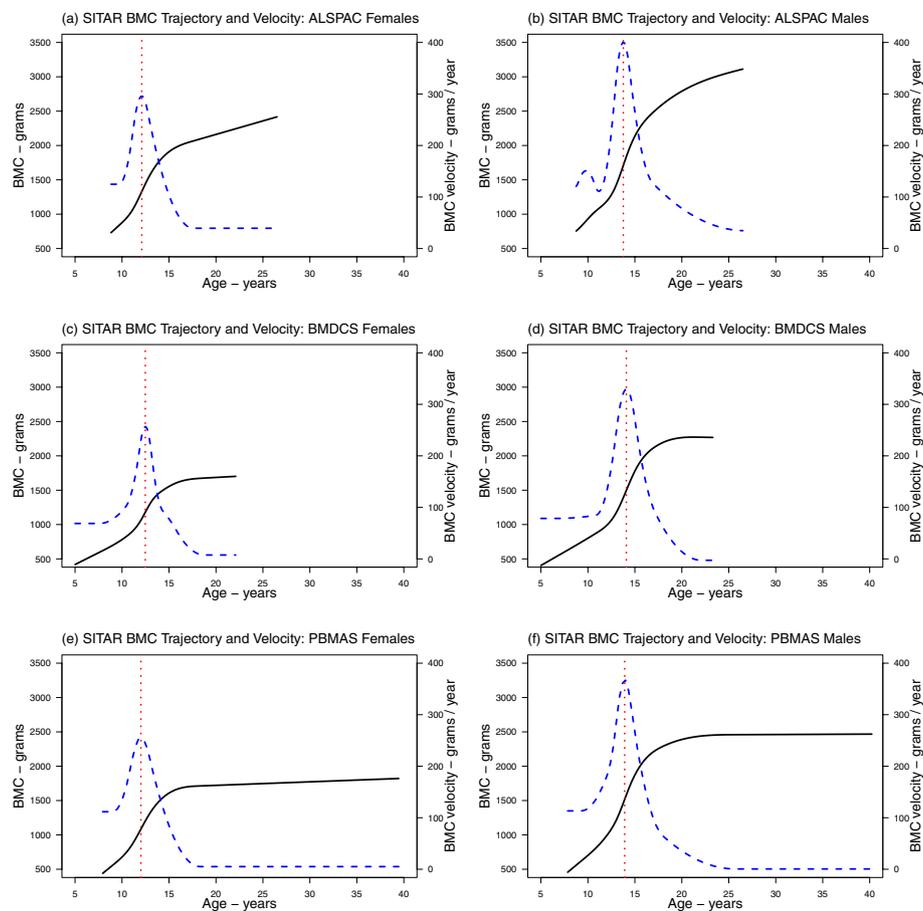
Our results provide evidence on bone mineral accrual from 5 to 40 years. Linear and natural cubic spline LME and SITAR showed that the levels and rates of change in BMC were greater for males than females, with peak gains in adolescence in both, but later in males than females. Growth mixture (latent trajectory) models (with trajectory shapes estimated using natural cubic splines) identified potentially distinct trajectory sub-groups, with the





greatest between-group differences seen in adolescence. While the main aim of these analyses was to illustrate different modelling approaches, our results are consistent

with previous studies on sex differences in BMC and suggest that in both sexes puberty is an important period for peak bone accrual [38–40].



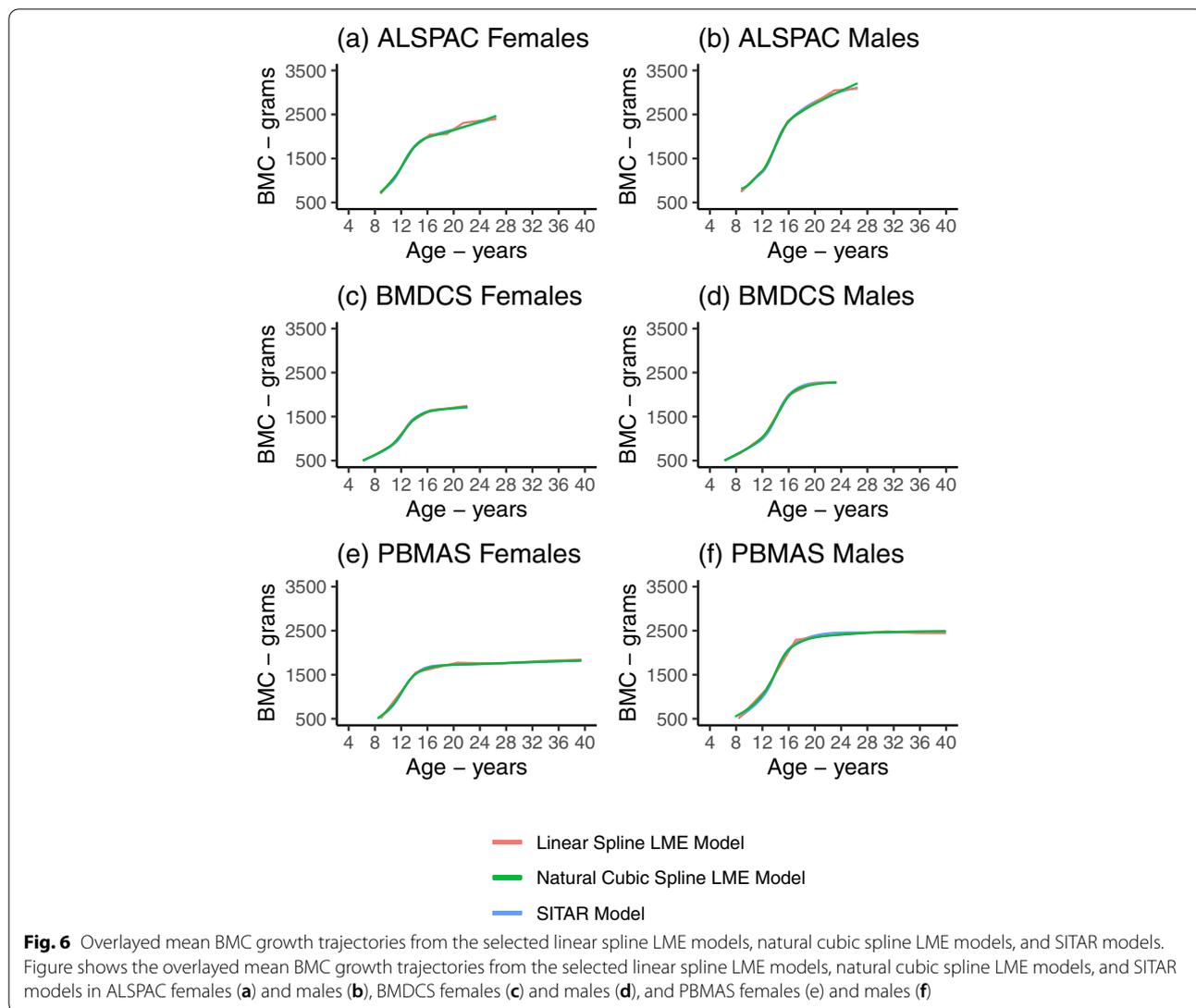
**Fig. 5** Mean BMC growth trajectory (solid black curves), mean BMC growth velocity (dashed blue curves), and mean age at peak BMC velocity (vertical red lines) from the selected SITAR models. Figure shows the estimated mean BMC trajectory (solid black curves), mean BMC growth velocity (dashed blue curves) and mean age at peak BMC velocity (vertical red lines) from SITAR models in ALSPAC females (a) and males (b), BMDCS females (c) and males (d), and PBMAS females (e) and males (f)

### Choosing a modelling approach

All the modelling techniques presented may be used to model BMC and other growth processes. Both spline LME models and SITAR were previously used on height/weight/BMI [4, 26, 41, 42]; linear splines LME and SITAR models were previously used on bone density [38, 43]; linear and natural cubic spline LME were used to analyse blood pressure change [44, 45]. Choice of method will be determined by research question, including whether this is concerned with differences in mean change over time in the whole study population (where LME or SITAR may be useful), or if the aim is to identify data-driven population subgroups for patterns of change (where a latent trajectory model may be useful). Complexity of the underlying trajectory, and data availability i.e., number of individuals and repeat measurements will also influence choice of method (see [46, 47] for a discussion on sample size in growth models). All models presented can handle unbalanced datasets (i.e., with individually varying

measurement occasions, as in our example). If the dataset is balanced (i.e., all measurements taken at the same schedule for everyone, e.g., participants are measured at exactly 2, 4, and 6 years of age), this will limit the complexity of models that can be fitted.

When the main aim is to quantify growth rate at different periods of the life-course, then linear splines may be preferred because of their more interpretable slope coefficients compared to the natural cubic spline LME and SITAR models. If the aim is to describe the shape of the trajectory or identify specific peaks and troughs (e.g., age at peak velocity), then natural cubic spline LME or SITAR are more useful, because linear spline cannot identify points of maxima/minima (but can identify periods/age windows). The number and spacing of repeated measures can influence model convergence and the complexity that can be allowed [48, 49]. SITAR was designed to model adolescent growth in height, and (like other nonlinear models) its parameters



reflect its specific purpose. Hence, SITAR may not work well for complex trajectories (e.g., depressive symptoms [50]), and natural cubic spline LME may offer more flexible alternatives. Notable, SITAR fitted without the timing fixed effect ( $\beta_0$ ) is analogous to a random intercept random slope model, and so should be at least as flexible as LME. A simple tactic that may improve SITAR (and optimise knot placement in LME models) is transforming the age scale [48, 51].

The regression splines presented have the advantage of being straightforward to estimate once the number/location of knots is set, with the functions calculated by solving linear equations. If the number of knots is too small, the spline can be sensitive to knot location (and placing knots at quantiles or evenly spaced may not be useful) whereas if the number of knots is too large this can lead to overfitting. Penalised regression splines (introduced

in the Methods) are highly flexible alternative functions which avoid the need for knot selection, though are less straightforward to estimate due to the uncertainty of the tuning parameter [12, 16, 17]. One powerful approach for applying penalised regression splines to describe population-average growth trajectories is using generalised additive mixed models (GAMM) [52–55]. Here, the linear predictor depends linearly on an unknown regression spline function of one or more covariates (e.g., time/age) and a tuning parameter is estimated (by REML, generalised cross-validation or other approaches) and used to penalise the spline and avoid overfitting. When applied to PBMAS, a penalised regression spline GAMM estimated mean BMC growth trajectory and velocity curves that were broadly consistent with the curves obtained from regression spline LME and SITAR models (Fig. 8).

**Table 3** Estimated mean BMC growth velocity during different age windows from the selected linear spline LME models

	Grams per year change in BMC [(mean (95% CI))]		
	ALSPAC	BMDCS	PBMAS
<i>Females</i>			
8.8y to 11.4y (ALSPAC); 5.0y to 7.9y (BMDCS); 7.9y to 14.2y (PBMAS)	177.7 (173.8 to 181.6)	67.6 (60.1 to 75.1)	195.4 (189.7 to 201.1)
11.4y to 13.9y (ALSPAC); 7.9y to 10.7y (BMDCS); 14.2y to 20.5y (PBMAS)	238.9 (236.0 to 241.7)	84.1 (79.0 to 89.2)	34.7 (30.0 to 39.5)
13.9y to 16.4y (ALSPAC); 10.7y to 13.6y (BMDCS); 20.5y to 26.9y (PBMAS)	117.3 (113.2 to 121.3)	189.0 (184.7 to 193.3)	-3.3 (-7.5 to 0.9)
16.4y to 18.9y (ALSPAC); 13.6y to 16.4y (BMDCS); 26.9y to 33.2y (PBMAS)	5.0 (-1.3 to 11.3)	85.7 (80.1 to 91.3)	9.1 (3.0 to 15.1)
18.9y to 21.5y (ALSPAC); 16.4y to 19.3y (BMDCS); 33.2y to 39.5y (PBMAS)	99.2 (74.2 to 124.2)	17.4 (10.6 to 24.1)	5.2 (-5.9 to 16.3)
21.5y to 24.0y (ALSPAC); 19.3y to 22.1y (BMDCS)	20.1 (-7.0 to 47.2)	19.5 (11.1 to 27.8)	-
24.0y to 26.5y (ALSPAC)	15.2 (3.0 to 27.5)	-	-
<i>Males</i>			
8.8y to 12.3y (ALSPAC); 5.0y to 8.6y (BMDCS); 7.8y to 12.5y (PBMAS)	148.5 (144.2 to 150.8)	70.2 (61.7 to 78.7)	157.5 (144.2 to 170.8)
12.3y to 15.9y (ALSPAC); 8.6y to 12.3y (BMDCS); 12.5y to 17.1y (PBMAS)	305.6 (302.5 to 308.8)	107.4 (101.7 to 113.1)	247.1 (238.7 to 255.4)
15.9y to 19.4y (ALSPAC); 12.3y to 16.0y (BMDCS); 17.1y to 21.7y (PBMAS)	103.3 (98.6 to 108.0)	253.2 (248.4 to 258.0)	19.3 (8.9 to 29.7)
19.4y to 23.0y (ALSPAC); 16.0y to 19.6y (BMDCS); 21.7y to 26.3y (PBMAS)	93.5 (84.1 to 102.8)	71.4 (63.3 to 79.5)	11.2 (0.8 to 21.6)
23.0y to 26.5y (ALSPAC); 19.6y to 23.3y (BMDCS); 26.3y to 31.0y (PBMAS)	11.7 (-6.4 to 29.9)	8.0 (-3.5 to 19.6)	10.7 (-1.5 to 22.8)
31.0y to 35.6y (PBMAS)	-	-	-7.9 (-26.4 to 10.7)
35.6y to 40.2y (PBMAS)	-	-	-0.4 (-32.7 to 32.0)

Age windows are defined by the number and position of the knots, which were placed at quantiles of the age distribution

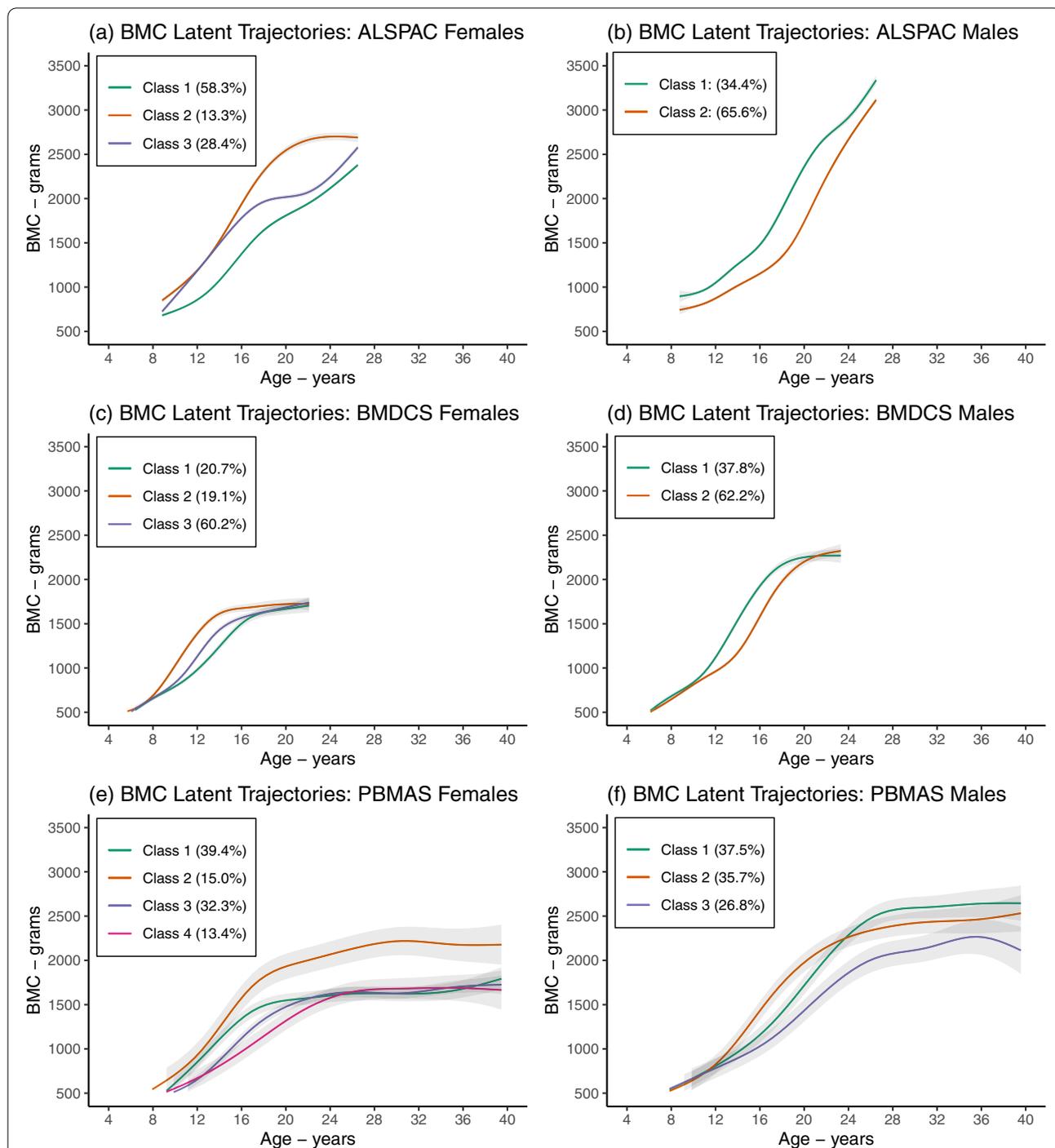
**Table 4** Estimated mean age at peak BMC velocity from the selected natural cubic spline LME models, and selected SITAR models. For comparison, the age windows for peak BMC velocity from the linear spline LME models are also presented

	Females			Males		
	ALSPAC	BMDCS	PBMAS	ALSPAC	BMDCS	PBMAS
<i>Mean age at peak BMC velocity (years)</i>						
Natural cubic spline LME model	12.5	12.4	12.2	13.9	14.4	14.1
SITAR model	12.1	12.5	12.0	13.8	14.1	13.9
<i>age window for peak BMC velocity from the linear spline LME model (years)</i>	11.4 to 13.9	10.7 to 13.6	7.9 to 14.2	12.3 to 15.9	12.3 to 16.0	12.5 to 17.1

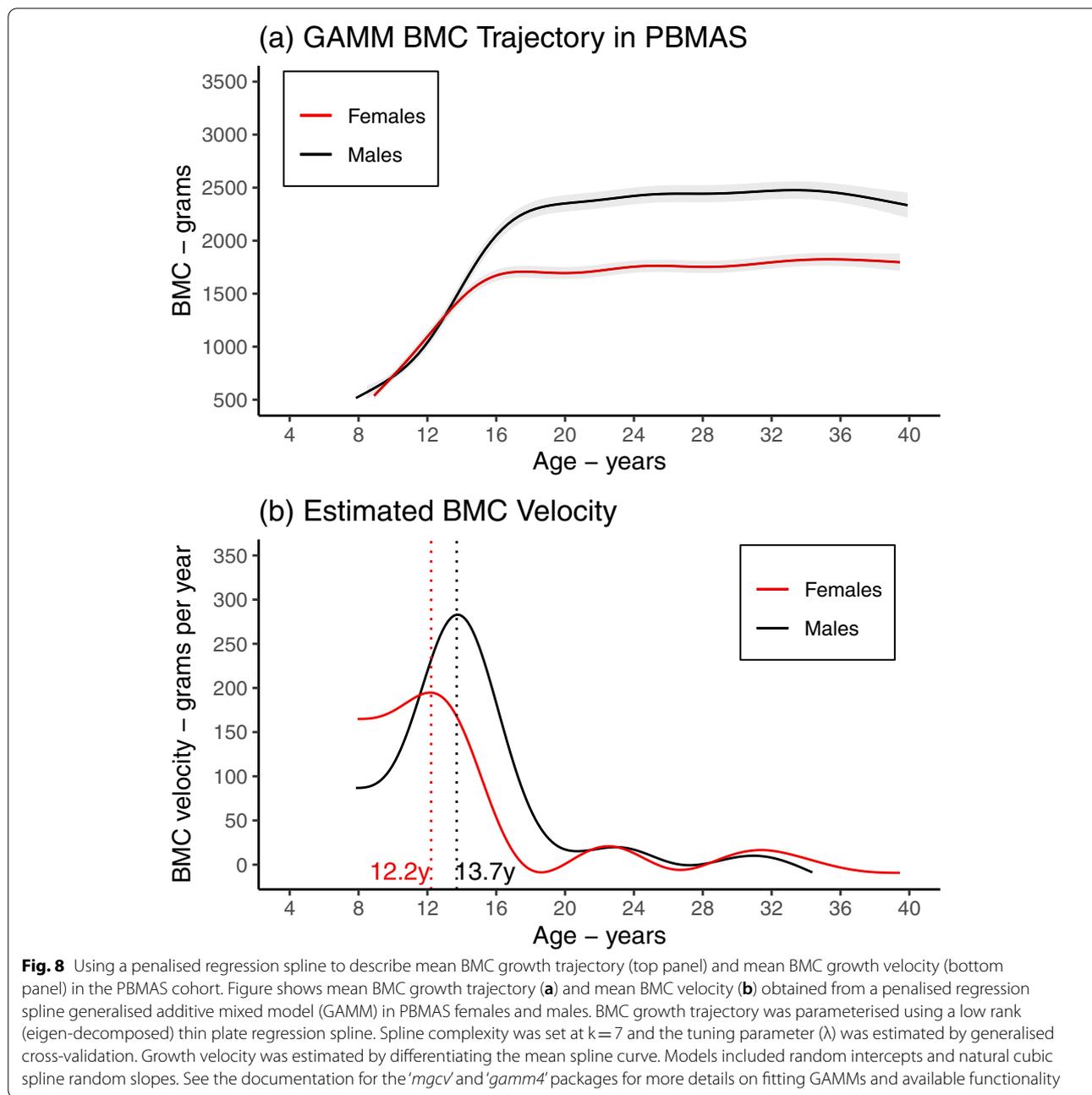
Distinct from the LME and SITAR models, we used growth mixture models with a natural cubic spline to identify population subgroups with different nonlinear BMC trajectories. Linear splines with estimated knots [56, 57] and smoothing splines [58] have also been used to parameterise the nonlinear curve in these models (i.e., instead of the natural cubic spline). Latent trajectory models (e.g., growth mixture models) have previously been used to identify trajectory subgroups for BMI [59], depressive symptoms [60], physical activity [61], glucose response [62], and environmental exposures [63], among others. It can be challenging to determine the optimal number of latent classes/sub-groups, including whether the classes are meaningful or if the model is just splitting the distribution of the random effects into a larger group and smaller extremes. Model selection is often subjective, and trajectory subgroups can be cohort-specific and may not replicate in other

cohorts therefore, it is important to follow established reporting guidelines [28].

Latent trajectory models will be most useful if there are strong reasons to hypothesise that a study population is composed of multiple unknown sub-groups of individuals with distinct trajectories. Less complex latent trajectory models (e.g., group-based trajectory models) are computationally more efficient than growth mixture models though, any subgroup differences identified by such models may just reflect within class variability, which is likely to be absorbed by random effects in growth mixture models. If the aim was in exploring specific hypotheses (e.g., sensitive periods), then clustering (data driven) approaches may not provide a suitable sub-group to test this, e.g., if the aim was to explore the hypothesis that a lower birth weight followed by faster growth in the first 1000 days increased cardiovascular risk, this approach might not identify



**Fig. 7** Mean BMC growth trajectories by subgroup (latent class) from the selected growth mixture models in ALSPAC females (a) and males (b), BMDCS females (c) and males (d), and PBMAS females (e) and males (f). Colours distinguish between latent trajectory subgroups within subplots and should not be used to compare between subplots. Shaded areas around the mean trajectories represent 95% confidence intervals. The numbers in each class are: ALSPAC females (class 1:  $n = 2337$ , class 2:  $n = 531$ , class 3:  $n = 1139$ ), ALSPAC males (class 1:  $n = 1339$ , class 2:  $n = 2549$ ), BMDCS females (class 1:  $n = 101$ , class 2:  $n = 93$ , class 3:  $n = 294$ ), BMDCS males (class 1:  $n = 176$ , class 2:  $n = 289$ ), PBMAS females (class 1:  $n = 50$ , class 2:  $n = 19$ , class 3:  $n = 41$ , class 4:  $n = 17$ ), PBMAS males (class 1:  $n = 42$ , class 2:  $n = 40$ , class 3:  $n = 30$ )

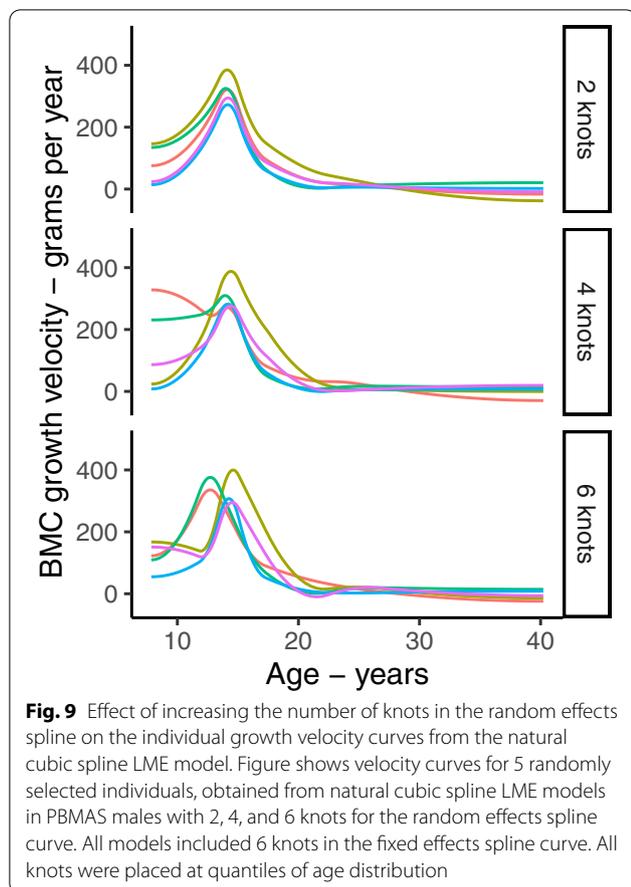


a cluster with this specific growth pattern. There may also be value in using a combination of approaches [64], like in this paper.

**Identifying determinants and outcomes of trajectories**

The models in this paper can be extended to include early life exposures and later outcomes to explore their associations with trajectories. Choice of method will depend on the research aims. Exposure variables can be

added within LME spline models as fixed effects and as interactions with splines to test their effect on trajectories and growth rate [3, 38, 65]. The individual growth features (e.g., peak velocity and age at peak velocity) can be obtained from spline LME model random effects and used in separate analyses as outcomes or exposures [18, 66] – however, it is important to allow enough complexity in the random-effects splines for sufficient between-person variability (Fig. 9). Individual growth features can



be easily obtained from SITAR and used in subsequent analyses to examine associations with exposures or outcomes [43, 67]. Of note, 2-stage approaches may be more biased than 1-stage joint model [68, 69]. Exposures and outcomes can be related to latent trajectories in a joint model or a multistage process where subgroups are first identified and subsequently used in separate models (unweighted or (preferably) weighed for classification probabilities) to examine associations [30, 59–63]. If the aim was to identify effects of repeated exposure, then a ‘structured’ modelling approach may be useful for testing competing hypotheses [70].

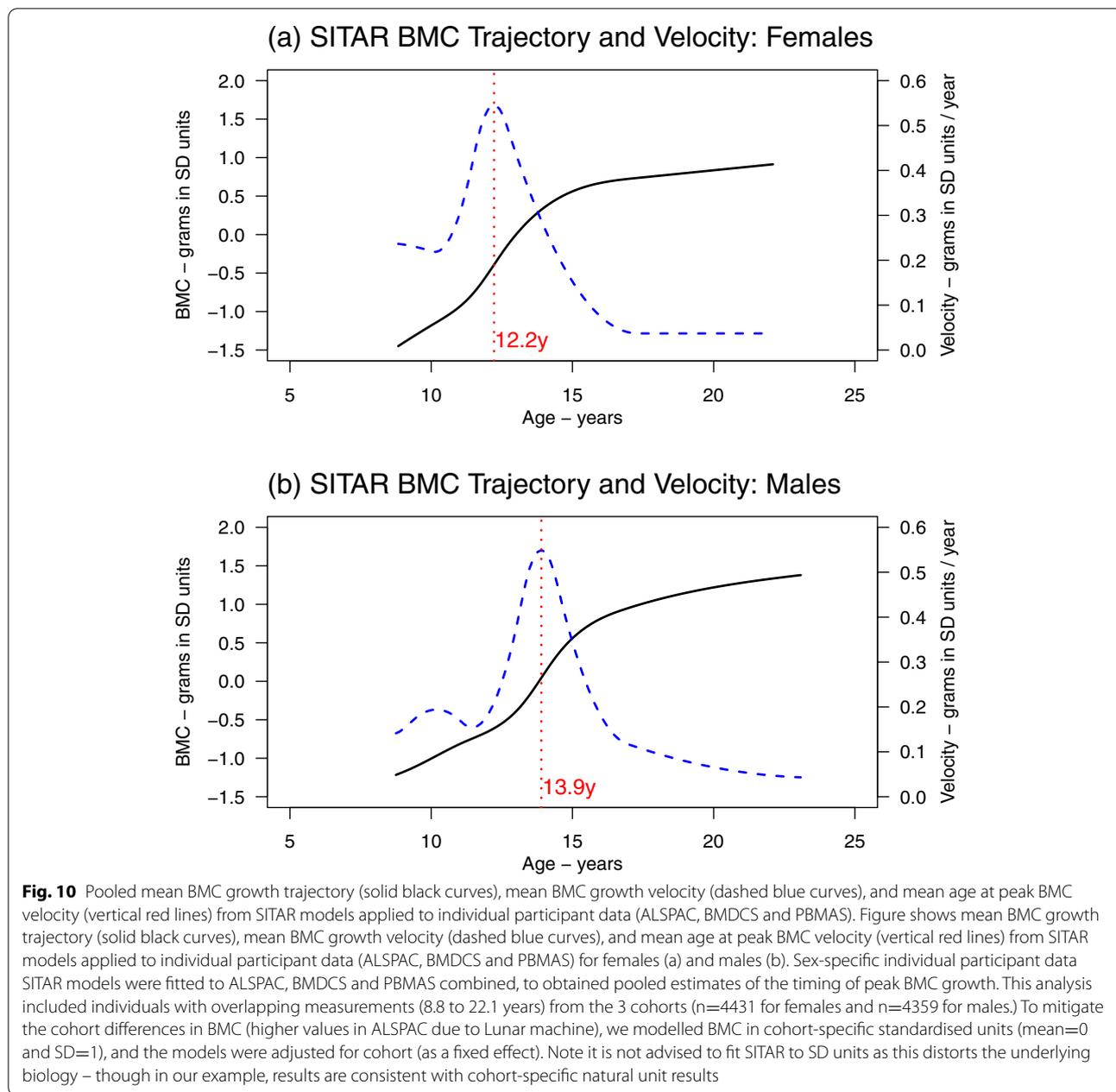
It is important to identify potential biases and explore ways for mitigating them when analysing (causal) associations between trajectories and exposures or outcomes in cohort studies. Missing data can bias associations depending on mechanism, and appropriate approaches to describe and handle missing data should be explored [12, 71–76]. In a repeated measure setting, individuals with missing outcome values can be included in the estimation sample if they have at least one observed outcome value. Mixed-effects models give unbiased results (i.e.,

not biased by missing data) when the probability that an outcome value is missing depends on observed values of the outcome (i.e., outcome missing at random (MAR) depending on observed outcome values). Bias due to missing data will occur for these models when the probability that the outcome is missing depends on underlying missing values (i.e., outcome missing not at random (MNAR) depending on missing outcome values). With incomplete covariates, bias can occur when the probability of excluding an individual with missing covariate data is related to the outcome. Regardless of bias, excluding individuals with missing covariate information will often mean discarding useful observed data, leading to imprecision [77].

Confounding can lead to spurious associations between trajectories and exposures or outcomes. Confounders (factors causally related to exposure and outcome) should be identified (e.g., using Directed Acyclic Graphs) and controlled for, taking care not to adjust for mediators (factors on the causal pathways) [78, 79]. Even with adjustment, residual confounding (from using poorly measured confounders or not adjusting for important confounders) can bias results. Some useful strategies for checking if residual confounding influences results include using negative control variables and comparing results across cohorts with different confounding structures [64, 80–82].

#### Comparing and modelling trajectories across cohorts

Researchers should be aware of potential differences in the participants, data collection methods and analysis models when comparing trajectories from different studies. For example, the higher BMC due to the Lunar machine in ALSPAC means it is inaccurate to conclude that Britons had higher BMC (and peak BMC velocity) than North Americans. Another example is the effect of medication use by older cohorts on combined blood pressure trajectories [83]. New and existing multicohort collaborations provide unique opportunities to jointly model such trajectories across different cohorts and to extend the amount of the life course studied [84, 85]. However, this also generates additional challenges including on model selection and missing data (see [86] for a discussion of challenges and solutions to multicohort modelling). Whatever approach is taken (cohort-specific or multicohort modelling), data harmonisation is an important initial step that involves making data comparable across studies [87, 88], e.g. using DXA reference standards to harmonise BMC [89]. Because age (but not BMC) was fully harmonised in our example, a simple approach to obtaining valid pooled estimate of age at peak BMC velocity is to fit a growth model to all



individuals with BMC expressed in cohort-specific SD units (Fig. 10).

**Conclusions**

LME models with linear and natural cubic splines, SITAR, and growth mixture models are useful for describing nonlinear growth trajectories in longitudinal population studies, and these methods can be adapted to other complex traits. Choice of method depends on research aims, complexity of the trajectory, and available

data. This illustrative paper and accompanying R analysis code and example datasets will we hope be a useful resource for researchers interested in modelling nonlinear longitudinal trajectories.

**Abbreviations**

ALSPAC: Avon Longitudinal Study of Parents and Children; BMC: Bone mineral content; BMDCS: Bone Mineral Density in Childhood Study; BMI: Body mass index; DXA: Dual-Energy X-ray Absorptiometry; GAMM: Generalised additive mixed model; GEE: Generalised estimating equations; LME: Linear mixed effects; MAR: Missing at random; ML: Maximum likelihood; MNAR: Missing not at random; PBMAS: Pediatric Bone Mineral Accrual Study; REML: Restricted maximum likelihood; SITAR: Super Imposition by Translation and Rotation.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-022-01542-8>.

**Additional file 1.** Cohort characteristics and participant numbers, and age and BMC at each visit.

**Additional file 2.** BIC and fit statistics for linear spline LME models, natural cubic spline LME models, and SITAR models with 2 to 6 knots in the fixed effects spline curve.

**Additional file 3.** Fit statistics and predicted mean BMC latent trajectories for growth mixture models with 1 to 5 latent classes.

## Acknowledgments

We thank the study participants from each cohort who took part in this study and the scientific and data collection teams from the ALSPAC, BMDCS and PBMAS cohorts. We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We acknowledge the BMDCS Principal Investigators: Vicente Gilsanz, MD, PhD, Heidi Kalkwarf, PhD, Joan Lappe, PhD, Sharon Oberfield, MD, John Shepherd, PhD, and the National Institute of Child Health and Human Development (NICHD) Data and Specimen Hub (DASH) for providing the BMDCS data that was used for this research.

## Authors' contributions

AE developed the idea for the paper with initial input from DAL and with further development from RAH, KMT, DLC, SAJ, TJC ASFK, ZL, SAF, ADGB, BSZ. AE conducted the analysis and wrote the first draft of the manuscript with initial input from DAL, RAH, KMT, TJC, and ZL provided advice on the description and interpretation of statistical methods. AE, RAH, KMT, DLC, SAJ, TJC ASFK, ZL, SAF, ADGB, BSZ, and DAL contributed to the development of the draft to its final version. ZL developed an R package to help complete this work (Package 'spluti' – Utility Functions for Post-Processing Univariate Interpolation Splines, Smoothing Splines and Regression Splines: <https://github.com/ZheyuanLi/spluti>). The author(s) read and approved the final manuscript.

## Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 874739 (LongITools) and No. 733206 (LifeCycle). AE, KMT, RAH and DAL work in a unit that is supported by the University of Bristol and UK Medical Research Council (MC\_UU\_00011/3 & MC\_UU\_00011/3). RAH is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 215408/Z/19/Z). ASFK is supported by an Economic Social Research Council Postdoctoral Fellowship (Grant ref.: ES/V011650/1). The UK Medical Research Council and Wellcome (Grant Ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grant funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). PBMAS was supported in part by funding from the Canadian Institutes of Health Research, the Saskatchewan Health Research Foundation (SHRF), The Dairy Farmers of Canada and the University of Saskatchewan. This research reflects only the authors' view, and the European Commission is not responsible for any use that may be made of the information it contains. The funders had no role in the design or conduct of the study, nor management, analysis, and interpretation of the data, or the decision to submit the manuscript for publication.

## Availability of data and materials

All the analysis code used in this paper can be found at <https://github.com/aelhak/nltmr/>. Details of all available data and the processes and procedures involved in accessing the ALSPAC resource can be found in the ALSPAC study website, which includes a fully searchable data dictionary and variable search tool (<http://www.bristol.ac.uk/alspac/researchers/our-data/>). The BMDCS data can be accessed through the NICHD DASH website (<https://dash.nichd.nih.gov/>). Researchers interested in accessing the PBMAS data should contact Professor Adam DG Baxter-Jones ([baxter.jones@usask.ca](mailto:baxter.jones@usask.ca)).

## Declarations

### Ethics approval and consent to participate

All cohorts had ethical approval from their relevant local or national ethics committees. The Avon Longitudinal Study of Parents and Children (ALSPAC) obtained ethics approval from the ALSPAC Law and Ethics committee and the local National Health Service Research Ethics Committee. The Bone Mineral Density in Childhood Study (BMDCS) study was approved by the Institutional Review Board of each clinical center: Children Hospital of Los Angeles (Los Angeles, CA), Cincinnati Children Hospital Medical Center (Cincinnati, OH), Creighton University (Omaha, NE), Children Hospital of Philadelphia (CHOP) (Philadelphia, PA), and Columbia University (New York, NY). The Pediatric Bone Mineral Accrual Study (PBMAS) study was approved by the University of Saskatchewan's biomedical review committee. All study participants gave written informed consent to participate in the respective cohorts and secondary data analyses.

### Consent for publication

All study participants gave written informed consent to participate in the respective cohorts and secondary data analyses.

### Competing interests

DAL reported grants from national and international government and charity funders, Roche Diagnostics, and Medtronic Ltd. for work unrelated to this publication. KMT reported grants from national and international government and charity funders, for work unrelated to this publication. No other disclosures were reported.

### Author details

<sup>1</sup>MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK. <sup>2</sup>Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. <sup>3</sup>Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. <sup>4</sup>Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA. <sup>5</sup>Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. <sup>6</sup>College of Kinesiology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada. <sup>7</sup>Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada. <sup>8</sup>UCL Great Ormond Street Institute of Child Health, London, UK. <sup>9</sup>Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK. <sup>10</sup>School of Mathematics and Statistics, Henan University, Kaifeng, Henan, China. <sup>11</sup>Department of Statistics and Actuarial Sciences, Simon Fraser University, Burnaby, BC, Canada. <sup>12</sup>Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. <sup>13</sup>Division of Endocrinology and Diabetes, Children's Hospital of Philadelphia, Philadelphia, PA, USA. <sup>14</sup>Division of Gastroenterology, Hepatology and Nutrition, Children's Hospital of Philadelphia, Philadelphia, PA, USA.

Received: 18 June 2021 Accepted: 11 February 2022

Published online: 15 March 2022

## References

1. Ben-Shlomo Y, Cooper R, Kuh D. The last two decades of life course epidemiology, and its relevance for research on ageing. *Int J Epidemiol*. 2016;45(4):973–88.
2. Grimm KJ, Ram N, Hamagami F. Nonlinear growth curves in developmental research. *Child Dev*. 2011;82(5):1357–71.
3. Howe LD, Tilling K, Matijasevich A, Petherick ES, Santos AC, Fairley L, et al. Linear spline multilevel models for summarising childhood growth trajectories: a guide to their application using examples from five birth cohorts. *Stat Methods Med Res*. 2016;25(5):1854–74.
4. Lourenço BH, Villamor E, Augusto RA, Cardoso MA. Influence of early life factors on body mass index trajectory during childhood: a population-based longitudinal analysis in the Western Brazilian Amazon. *Matern Child Nutr*. 2015;11(2):240–52.
5. Cole TJ, Donaldson MD, Ben-Shlomo Y. SITAR—a useful instrument for growth curve analysis. *Int J Epidemiol*. 2010;39(6):1558–66.

6. Herle M, Micali N, Abdulkadir M, Loos R, Bryant-Waugh R, Hübel C, et al. Identifying typical trajectories in longitudinal data: modelling strategies and interpretations. *Eur J Epidemiol*. 2020;35(3):205–22.
7. Sauerbrei W, Abrahamowicz M, Altman DG, le Cessie S, Carpenter J, initiative obotS. STREngthening analytical thinking for observational studies: the STRATOS initiative. *Stat Med*. 2014;33(30):5413–32.
8. Tu YK, Tilling K, Sterne JA, Gilthorpe MS. A critical evaluation of statistical approaches to examining the role of growth trajectories in the developmental origins of health and disease. *Int J Epidemiol*. 2013;42(5):1327–39.
9. Curran PJ, Obeidat K, Losardo D. Twelve frequently asked questions about growth curve modeling. *J Cogn Dev*. 2010;11(2):121–36.
10. Macdonald-Wallis C, Lawlor DA, Palmer T, Tilling K. Multivariate multilevel spline models for parallel growth processes: application to weight and mean arterial pressure in pregnancy. *Stat Med*. 2012;31(26):3147–64.
11. Twisk JW. Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis. *Eur J Epidemiol*. 2004;19(8):769–76.
12. Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G. Longitudinal data analysis. USA: Chapman & Hall/CRC; 2009.
13. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982;38(4):963–74.
14. Goldstein H, De Stavola B. Statistical modelling of repeated measurement data. *Longitud Life Course Stud*. 2010;1(2):170–85.
15. Cole TJ: sitar: Super Imposition by Translation and Rotation growth curve analysis. R package version 1.2.0. 2021. <https://cran.r-project.org/web/packages/sitar/index.html>.
16. Perperoglou A, Sauerbrei W, Abrahamowicz M, Schmid M. A review of spline function procedures in R. *BMC Med Res Methodol*. 2019;19(1):46.
17. Suk HW, West SG, Fine KL, Grimm KJ. Nonlinear growth curve modeling using penalized spline models: a gentle introduction. *Psychol Methods*. 2019;24(3):269–90.
18. Aris IM, Bernard JY, Chen LW, Tint MT, Pang WW, Lim WY, et al. Infant body mass index peak and early childhood cardio-metabolic risk markers in a multi-ethnic Asian birth cohort. *Int J Epidemiol*. 2017;46(2):513–25.
19. Fonseca MJ, Moreira C, Santos AC. Adiposity rebound and cardiometabolic health in childhood: results from the generation XXI birth cohort. *Int J Epidemiol*. 2021;50(4):1260–71.
20. Harrell F. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. 1st ed. New York: Springer; 2001.
21. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med*. 2010;29(9):1037–57.
22. Mackenzie ML, Donovan CR, McArdle BH. Regression spline mixed models: A forestry example. *J Agric Biol Environ Stat*. 2005;10(4):394.
23. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. New York: Springer; 2017.
24. Naumova EN, Must A, Laird NM. Tutorial in biostatistics: evaluating the impact of ‘critical periods’ in longitudinal studies of growth using piecewise mixed effects models. *Int J Epidemiol*. 2001;30(6):1332–41.
25. Beath KJ. Infant growth modelling using a shape invariant model with random effects. *Stat Med*. 2007;26(12):2547–64.
26. Cole TJ, Kuh D, Johnson W, Ward KA, Howe LD, Adams JE, et al. Using super-imposition by translation and rotation (SITAR) to relate pubertal growth to bone health in later life: the Medical Research Council (MRC) National Survey of health and development. *Int J Epidemiol*. 2016.
27. Berlin KS, Parra GR, Williams NA. An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. *J Pediatr Psychol*. 2013;39(2):188–203.
28. van de Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GROLTS-checklist: guidelines for reporting on latent trajectory studies. *Struct Equ Model Multidiscip J*. 2017;24(3):451–67.
29. Lennon H, Kelly S, Sperrin M, Buchan I, Cross AJ, Leitzmann M, et al. Framework to construct and interpret latent class trajectory modelling. *BMJ Open*. 2018;8(7):e020683.
30. Proust-Lima C, Philipps V, Liqueur B. Estimation of extended mixed models using latent classes and latent processes: the R package lcm. *J Stat Softw*. 2017;1(Issue 2):2017.
31. Harvey N, Dennison E, Cooper C. Osteoporosis: a lifecourse approach. *J Bone Miner Res*. 2014;29(9):1917–25.
32. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort profile: the Avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42(1):97–110.
33. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: the ‘children of the 90s’-the index offspring of the Avon longitudinal study of parents and children. *Int J Epidemiol*. 2013;42(1):111–27.
34. McCormack SE, Cousminer DL, Chesi A, Mitchell JA, Roy SM, Kalkwarf HJ, et al. Association between linear growth and bone accrual in a diverse cohort of children and adolescents. *JAMA Pediatr*. 2017;171(9):e171769.
35. Baxter-Jones AD, Faulkner RA, Forwood MR, Mirwald RL, Bailey DA. Bone mineral accrual from 8 to 30 years of age: an estimation of peak bone mass. *J Bone Miner Res*. 2011;26(8):1729–39.
36. Nowok B, Raab GM, Dibben C. synthpop: bespoke creation of synthetic data in R. *J Stat Softw*. 2016;74(11):26.
37. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software*. 2015;67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>.
38. Elhakeem A, Fryszyk M, Tilling K, Tobias JH, Lawlor DA. Association between age at puberty and bone accrual from 10 to 25 years of age. *JAMA Netw Open*. 2019;2(8):e198918.
39. Jackowski SA, Erlandson MC, Mirwald RL, Faulkner RA, Bailey DA, Kontulainen SA, et al. Effect of maturational timing on bone mineral content accrual from childhood to adulthood: evidence from 15 years of longitudinal data. *Bone*. 2011;48(5):1178–85.
40. Cousminer DL, Mitchell JA, Chesi A, Roy SM, Kalkwarf HJ, Lappe JM, et al. Genetically determined later puberty impacts lowered bone mineral density in childhood and adulthood. *J Bone Miner Res*. 2018;33(3):430–6.
41. Khera AV, Chaffin M, Wade KH, Zahid S, Brancale J, Xia R, et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell*. 2019;177(3):587–596.e589.
42. Jensen SM, Ritz C, Ejlerskov KT, Mølgaard C, Michaelsen KF. Infant BMI peak, breastfeeding, and body composition at age 3 y. *Am J Clin Nutr*. 2014;101(2):319–25.
43. Cousminer DL, Wagley Y, Pippin JA, Elhakeem A, Way GP, Pahl MC, et al. Genome-wide association study implicates novel loci and reveals candidate effector genes for longitudinal pediatric bone accrual. *Genome Biol*. 2021;22(1):1.
44. O’Keefe LM, Simpkin AJ, Tilling K, Anderson EL, Hughes AD, Lawlor DA, et al. Sex-specific trajectories of measures of cardiovascular health during childhood and adolescence: a prospective cohort study. *Atherosclerosis*. 2018;278:190–6.
45. Lambert PC, Abrams KR, Jones DR, Halligan AW, Shennan A. Analysis of ambulatory blood pressure monitor data using a hierarchical model incorporating restricted cubic splines and heterogeneous within-subject variances. *Stat Med*. 2001;20(24):3789–805.
46. Snijders T. Power and sample size in multilevel modeling. In: Everitt BS, Howell DC, editors. *Encyclopedia of Statistics in Behavioral Science*. Chichester: Wiley; 2005.
47. Guo Y, Logan HL, Glueck DH, Muller KE. Selecting a sample size for studies with repeated measures. *BMC Med Res Methodol*. 2013;13(1):100.
48. Simpkin AJ, Sayers A, Gilthorpe MS, Heron J, Tilling K. Modelling height in adolescence: a comparison of methods for estimating the age at peak height velocity. *Ann Hum Biol*. 2017;44(8):715–22.
49. Tilling K, Macdonald-Wallis C, Lawlor DA, Hughes RA, Howe LD. Modelling childhood growth using fractional polynomials and linear splines. *Ann Nutr Metab*. 2014;65(2–3):129–38.
50. Kwong ASF, Manley D, Timpson NJ, Pearson RM, Heron J, Sallis H, et al. Identifying critical points of trajectories of depressive symptoms from childhood to young adulthood. *J Youth Adolesc*. 2019;48(4):815–27.
51. Cole TJ. Optimal design for longitudinal studies to estimate pubertal height growth in individuals. *Ann Hum Biol*. 2018;45(4):314–20.
52. Wood SN. Generalized additive models an introduction with R. 2nd ed: Chapman & Hall/CRC; USA: 2017.
53. Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ*. 2019;7:e6876.
54. Wood SN. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. 2021. p. 1.8–136 <https://cran.r-project.org/web/packages/mgcv/index.html>.

55. Wood SN, Scheipl F. *gamm4: generalized additive mixed models using 'mgcv' and 'lme4'*; 2017. p. 0.2–5. <http://cran.r-project.org/web/packages/gamm4/index.html>
56. Kohli N, Haring JR, Zopluoglu C. A finite mixture of nonlinear random coefficient models for continuous repeated measures data. *Psychometrika*. 2016;81(3):851–80.
57. Lock EF, Kohli N, Bose M. Detecting multiple random changepoints in Bayesian piecewise growth mixture models. *Psychometrika*. 2018;83(3):733–50.
58. Ding M, Chavarro JE, Fitzmaurice GM. Development of a mixture model allowing for smoothing functions of longitudinal trajectories. *Stat Methods Med Res*. 2021;30(2):549–62.
59. Buscot M-J, Thomson RJ, Juonala M, Sabin MA, Burgner DP, Lehtimäki T, et al. Distinct child-to-adult body mass index trajectories are associated with different levels of adult cardiometabolic risk. *Eur Heart J*. 2018;39(24):2263–70.
60. Kwong ASF, Lopez-Lopez JA, Hammerton G, Manley D, Timpson NJ, Leckie G, et al. Genetic and environmental risk factors associated with trajectories of depression symptoms from adolescence to young adulthood. *JAMA Netw Open*. 2019;2(6):e196587.
61. Elhakeem A, Heron J, Tobias JH, Lawlor DA. Physical activity throughout adolescence and peak hip strength in young adults. *JAMA Netw Open*. 2020;3(8):e2013463.
62. Hulman A, Witte DR, Vistisen D, Balkau B, Dekker JM, Herder C, et al. Pathophysiological characteristics underlying different glucose response curves: a latent class trajectory analysis from the prospective EGIR-RISC study. *Diabetes Care*. 2018;41(8):1740–8.
63. Lévêque E, Lacourt A, Philipps V, Luce D, Guénel P, Stücker I, et al. A new trajectory approach for investigating the association between an environmental or occupational exposure over lifetime and the risk of chronic disease: application to smoking, asbestos, and lung cancer. *Plos One*. 2020;15(8):e0236736.
64. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45(6):1866–86.
65. Madden JM, Li X, Kearney PM, Tilling K, Fitzgerald AP. Exploring diurnal variation using piecewise linear splines: an example using blood pressure. *Emerg Themes Epidemiol*. 2017;14:1–1.
66. Brilleman SL, Howe LD, Wolfe R, Tilling K. Bayesian piecewise linear mixed models with a random change point: an application to BMI rebound in childhood. *Epidemiology*. 2017;28(6):827–33.
67. Crozier SR, Johnson W, Cole TJ, Macdonald-Wallis C, Muniz-Terrera G, Inskip HM, et al. A discussion of statistical methods to characterise early growth and its impact on bone mineral content later in childhood. *Ann Hum Biol*. 2019;46(1):17–26.
68. Sayers A, Heron J, Smith A, Macdonald-Wallis C, Gilthorpe MS, Steele F, et al. Joint modelling compared with two stage methods for analysing longitudinal data and prospective outcomes: a simulation study of childhood growth and BP. *Stat Methods Med Res*. 2017;26(1):437–52.
69. Parker RMA, Leckie G, Goldstein H, Howe LD, Heron J, Hughes AD, et al. Joint modeling of individual trajectories, within-individual variability, and a later outcome: systolic blood pressure through childhood and left ventricular mass in early adulthood. *Am J Epidemiol*. 2021;190(4):652–62.
70. Smith AD, Hardy R, Heron J, Joinson CJ, Lawlor DA, Macdonald-Wallis C, et al. A structured approach to hypotheses involving continuous exposures over the life course. *Int J Epidemiol*. 2016;45(4):1271–9.
71. Lee KJ, Tilling K, Cornish RP, Little RJ, Bell ML, Goetghebuer E, et al. Framework for the treatment and reporting of missing data in observational studies: the TARMOS framework. *J Clin Epidemiol*. 2021;134:79–88.
72. van Buuren S. *Flexible imputation of missing data*. 2nd ed. Chapman & Hall/CRC. USA; 2018.
73. Matteo Quartagno SG, Carpenter J. *jomo: a flexible package for two-level joint modelling multiple imputation*. *R J*. 2019;11(2):205–28.
74. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *Int J Epidemiol*. 2019;48(4):1294–304.
75. Twisk J, de Boer M, de Vente W, Heymans M. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. *J Clin Epidemiol*. 2013;66(9):1022–8.
76. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Med Res Methodol*. 2018;18(1):168.
77. Huque MH, Moreno-Betancur M, Quartagno M, Simpson JA, Carlin JB, Lee KJ. Multiple imputation methods for handling incomplete longitudinal and clustered data where the target analysis is a linear mixed effects model. *Biom J*. 2020;62(2):444–66.
78. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34(3):211–9.
79. Groenwold RHH, Palmer TM, Tilling K. To Adjust or Not to Adjust? When a "Confounder" Is Only Measured After Exposure. *Epidemiology*. 2021;32(2):194–201. <https://doi.org/10.1097/EDE.0000000000001312>.
80. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383–8.
81. Taylor K, Elhakeem A, Nader JLT, Yang T, Isaevska E, Richiardi L, et al. Effect of maternal prepregnancy/early-pregnancy BMI and pregnancy smoking and alcohol on congenital heart diseases: a parental negative control study. *J Am Heart Assoc*. 2021;10(11):e020051.
82. Brion MJ, Lawlor DA, Matijasevich A, Horta B, Anselmi L, Araújo CL, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*. 2011;40(3):670–80.
83. Wills AK, Lawlor DA, Matthews FE, Aihie Sayer A, Bakra E, Ben-Shlomo Y, et al. Life course trajectories of systolic blood pressure using longitudinal data from eight UK cohorts. *PLoS Med*. 2011;8(6):e1000440.
84. Jaddoe VVW, Felix JF, Andersen AN, Charles MA, Chatzi L, Corpeleijn E, et al. The LifeCycle project-EU child cohort network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. *Eur J Epidemiol*. 2020;35(7):709–24.
85. Ronkainen J, Nedelec R, Atehortua A, Balkhiyarova Z, Zhanna A, Dang V, et al. LongTools: dynamic longitudinal exposome trajectories in cardiovascular and metabolic non-communicable diseases. *Environ Epidemiol*. 2021;6(1):e184. <https://doi.org/10.1097/EE9.0000000000000184>.
86. Hughes RA, Tilling K, Lawlor DA. Combining longitudinal data from different cohorts to examine the life-course trajectory. *Am J Epidemiol*. 2021;190(12):2680–9.
87. Pinot de Moira A, Haakma S, Strandberg-Larsen K, van Enckevort E, Kooijman M, Cadman T, et al. The EU Child Cohort Network's core data: establishing a set of findable, accessible, interoperable and re-usable (FAIR) variables. *Eur J Epidemiol*. 2021;36(5):565–80.
88. Nader JL, López M, Julvez J, Guxens M, Cadman T, Elhakeem A, et al. Cohort description: measures of early-life behaviour and later psychopathology in the LifeCycle project - EU child cohort network. *J Epidemiol*. 2021. (Epub ahead of print). <https://doi.org/10.2188/jea.JE20210241>.
89. Baxter-Jones AD, Burrows M, Bachrach LK, Lloyd T, Petit M, Macdonald H, et al. International longitudinal pediatric reference standards for bone mineral content. *Bone*. 2010;46(1):208–16.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

