Journal of Data Science 22 (1), 116–137
January 2024
DOI: 10.6339/23-JDS1104
Data Science in Action

Association Between Body Fat and Body Mass Index from Incomplete Longitudinal Proportion Data: Findings from the Fels Study

Xin Tong[1], Seohyun Kim[2], Dipankar Bandyopadhyay[3,*], and Shumei Sun[3]

[1]*Dept. of Psychology, University of Virginia, Charlottesville, VA, USA*
[2]*Mid-Atlantic Permanente Research Institute, Kaiser Permanente, Rockville, MD, USA*
[3]*Dept. of Biostatistics, Virginia Commonwealth University, Richmond, VA, USA*

Abstract

Obesity rates continue to exhibit an upward trajectory, particularly in the US, and is the underlying cause of several comorbidities, including but not limited to high blood pressure, high cholesterol, diabetes, heart disease, stroke, and cancers. To monitor obesity, body mass index (BMI) and proportion body fat (PBF) are two commonly used measurements. Although BMI and PBF changes over time in an individual's lifespan and their relationship may also change dynamically, existing work has mostly remained cross-sectional, or separately modeling BMI and PBF. A combined longitudinal assessment is expected to be more effective in unravelling their complex interplay. To mitigate this, we consider Bayesian cross-domain latent growth curve models within a structural equation modeling framework, which simultaneously handles issues such as individually varying time metrics, proportion data, and potential missing not at random data for joint assessment of the longitudinal changes of BMI and PBF. Through simulation studies, we observe that our proposed models and estimation method yielded parameter estimates with small bias and mean squared error in general, however, a mis-specified missing data mechanism may cause inaccurate and inefficient parameter estimates. Furthermore, we demonstrate application of our method to a motivating longitudinal obesity study, controlling for both time-invariant (such as, sex), and time-varying (such as diastolic and systolic blood pressure, biceps skinfold, bioelectrical impedance, and waist circumference) covariates in separate models. Under time-invariance, we observe that the initial BMI level and the rate of change in BMI influenced PBF. However, in presence of time-varying covariates, only the initial BMI level influenced the initial PBF. The added-on selection model estimation indicated that observations with higher PBF values were less likely to be missing.

Keywords *cross-domain latent growth curve model; individually varying time metrics; missing data; obesity; proportion data*

1 Introduction

Obesity is a growing problem worldwide. Although at its simplest, obesity results from people consuming more calories than their bodies can burn, and seems preventable by changes in dietary and physical activity patterns. However, the solution is much more complicated than "eating

*Corresponding author: Dept. of Biostatistics, Virgina Commonwealth University, 830 E. Main Street, PO Box 980032, Richmond, VA, 23298-0032. Email: bandyopd@gmail.com.

Received May 20, 2022; Accepted May 14, 2023

less and exercising more". Current data indicate that the best diet (e.g., low-fat, or low carb) varies across individuals, and exercise only helps 3% to 5% individuals with losing weight and may introduce possible adverse effects (e.g., Makris and Foster, 2011; McQueen, 2009; Petridou et al., 2019). Over the past 45 years, obesity rate for adults in the United States (U.S.) has tripled from 12% to 36%; see World Health Organization report (WHO, 2017). Obesity is the underlying cause of many diseases and comorbidities, including but not limited to high blood pressure, high cholesterol, diabetes, heart disease, stroke, depresseion, anxiety, and even cancer. Recent studies have also shown that obesity is associated with severe forms of the Coronavirus Disease 2019 (COVID-19; e.g., Caussy et al., 2020; Samuels, 2020), and is second only to older age as the main driver for people needing hospitalization due to COVID-19. The estimated annual medical cost burden of obesity was $147 billion U.S. dollars (in 2018), with $1,429 higher per-capita cost for the obese, compared to the normal weight.

To assess and monitor obesity, two important and popularly used biomarkers are the body mass index (BMI), and proportion body fat (PBF), which also appear to be highly correlated (Jelena et al., 2016; Nasr Eldeen et al., 2017). BMI is a numerical (continuous) measure, defined as the ratio of the weight (in kilograms) to the square of the height (in meters^2). It does not distinguish between lean muscle and fat mass, often leading to confusion while categorizing subjects who are fit with higher body weight, or who are not fit with lower weight. BMI also does not account for gender, age, and ethnicity, and hence may not be equally valid for different populations. On the other hand, PBF, defined as the ratio of the total fat mass to the total body mass, is a proportion $\in (0, 1)$, and the only body measurement which directly calculates a subject's relative body composition, without regards to height or weight. However, methods used to quantify PBF are often expensive, or inaccurate (Wells and Fewtrell, 2006).

The relationship between BMI and PBF for adults has been earlier explored in various studies (Deurenberg et al., 1991; Gallagher et al., 1996; Jackson et al., 2002) via linear prediction formulas, controlling for age and sex/gender. For example, Deurenberg et al. (1991) proposed the widely accepted empirical formula: $\text{PBF} = (1.20 \times \text{BMI}) + (0.23 \times \text{Age}) - (10.8 \times \text{sex}) - 5.4$, where Age is in years, and the binary indicator of sex is 1(0) for males (females). However, these formulas have important limitations. For example, the Deurenberg et al. (1991) formula was based on a cross-sectional study, where the effect of weight gain (and hence BMI) over time may not be efficiently captured. Also, these formulas assumed a linear relationship between BMI and PBF, whereas, controversy exists as to whether the relationship is in fact curvilinear/quadratic (Dulloo et al., 2010; Ranasinghe et al., 2013; Ho-Pham et al., 2015). Moreover, the relationship between BMI and PBF can be dynamic and time-varying, and thus, even a quadratic form may not provide an accurate representation of the complex relationship.

Since obesity is a chronic disease and cannot be overcome quickly, longitudinal/repeated measure studies appear as effective tools to help unravel the complex disease evolution over time as captured by plausible endpoints (here, BMI and PBF), and its relationship to other factors (covariates), factoring in within-subject association, and between-subject variations. In that vein, Demerath et al. (2004) examined the effects of birth cohort and rate of maturation on the timing and pattern of increases in BMI during adolescence in girls. The motivation of this paper comes from the Fels (Roche, 1992) longitudinal study, which has been earlier considered (Demerath et al., 2006; Guo et al., 1997, 1999; Sabo et al., 2012b; Sun et al., 2012) in the context of modeling BMI and PBF. For example, Guo et al. (1997) investigated patterns of PBF changes for 244 eight to twenty-year-old Caucasians using a cubic model with random intercepts and slopes, and found that the pattern depended on gender and rate of maturation. Also, Guo et al. (1999) fitted linear mixed-effects models for BMI and PBF, and found that physical activity was

associated with reduction in BMI and PBF for middle-aged men and increases in PBF among women. Furthermore, Sun et al. (2012) examined secular trends (by birth decade) in BMI and PBF for children and adolescents, separately, with polynomials of various orders specified to model possible non-linear relationship of age. However, the above studies precludes modeling the simultaneous complex evolution of the changes in BMI and PBF.

For a comprehensive understanding of the dynamics of BMI and PBF evolution, longitudinal structural equation modeling, or SEM (Little, 2013) is often the method of choice. SEM is a popular multivariate modeling technique in social, behavioral, and medical sciences, allowing for testing complex theories by modeling manifest variables, latent variables and measurement errors, simultaneously. Longitudinal SEM, in particular, considers multivariate modeling of changes over time, and is used to answer lifespan relevant questions, such as the evaluation of the average trajectory of endpoints, the variability between subjects, and what predicts this variability. In this context, latent growth curve models (LGCM), based on the SEM framework, is a powerful set of techniques that analyzes trends over time and the variations in changes over time for study subjects, in terms of an underlying, latent, unobserved process (Duncan and Duncan, 2004).

In this paper, we propose to use cross-domain LGCMs (Lee and Whittaker, 2018) to investigate how the changes in BMI affects the changes in PBF over time, in light of evaluating obesity from the motivating Fels data (Roche, 1992). Specifically, we investigate factors such as gender, waist circumference, blood pressure, etc (and their meaningful interpretation), that influence the changes in BMI and PBF using longitudinal SEMs. The data presents various technical challenges, such as modeling proportion data, handling individually varying time metrics, and missing data issues. The key contributions in this paper includes: (a) combining the changes of BMI and PBF via one longitudinal SEM model and studying their interrelationships, (b) handling non-ignorable missing data, i.e., missing-not-at-random, or MNAR (Enders, 2011) within the cross-domain latent growth curve modeling, (c) introducing an added-on beta distribution structure to the cross-domain LGCM to model continuous proportion data, and (d) using a definition variable approach to account for the individually varying time metrics within the SEM framework. For inference, we consider a Bayesian route powered by Markov chain Monte Carlo (MCMC) techniques that incorporate prior information on model parameters, naturally accommodates missing data specification, and factors in seamless uncertainty quantification at every step of the model hierarchy to yield posterior parameter estimates within complex model structures.

The rest of this paper is organized as follows. After a brief introduction to the Fels data in Section 2, the statistical modeling outlined in Section 3 proposes the cross-domain LGCMs to assess the changes in BMI and PBF domains simultaneously with a time invariant covariate (in Study 1), and additional multiple time-varying covariates (in Study 2). Utilizing a Bayesian paradigm, the proportion PBF responses were modeled as a beta density, with individually-varying time metrics in the SEM handled via a definition variable approach, accounting for missing data using added-on selection models. Next, Section 4 presents the results from the model fits to the two studies. The finite-sample performances of our cross-domain LGCMs along with robustness assessments in regards to missing data assumptions were evaluated in Section 5. Finally, Section 6 concludes, pointing to plausible future research.

2 Fels Data

Our motivating data in this paper was obtained from participants of the Fels longitudinal study, the world's largest and earliest longitudinal study of human growth and body composition (Roche, 1992). Since 1929, more than 1600 subjects have been enrolled (predominantly Caucasian), with health characteristics measured at birth, 1, 3, 6, 9, and 12 months, then semi-annually to 18 years, and biennially thereafter. Our current analytical sample consists of data collected on and after August 1976 (when PBF was actually measured) till June 2010 for participants with at least two responses (for both BMI and PBF), which reduced our analytical sample to 431 female and 395 male subjects, with 4876 total observations. Subjects were measured between 2 to 15 times (average = 5.9; standard deviation = 3.25). Across all observations, the mean age is 29.52 with a minimum of 6.004 and a maximum of 84.827. Along with BMI and PBF, various other health status indicators are available, which includes diastolic blood pressure (DBP), systolic blood pressure (SBP), biceps skinfold (Bicep), bioelectrical impedance (BCimp), and waist circumference (waist). The PBF variable has 175 missing observations.

As discussed in Section 1, the change patterns of BMI and PBF from cross-sectional studies were not consistent (can be linear, or curvillinear). Although there exist studies analyzing the longitudinal trajectories of BMI and PBF, separately, there is no research investigating the longitudinal relationship between their change patterns. The Fels data allow us to perform a systematic longitudinal study to investigate such relationship, and potential factors that may influence the relationship.

3 Statistical Models

In this section, we outline our statistical modeling in light of the aforementioned two studies. Based on published literature (Deurenberg et al., 1991; Jackson et al., 2002) which suggests BMI and PBF trajectories are dependent on sex and age, we consider sex as a time-invariant covariate, and age as the time variable in our cross-domain LGCM in Study 1. Also, following previous research (Sabo et al., 2012b; Ho-Pham et al., 2015) which suggests considering the effects of other factors, we modify Study 1 to further control for additional available time-varying covariates, such as DBP, SBP, Bicep, BCimp, and waist, to investigate the changes of BMI and PBF over time in Study 2.

3.1 Study 1: Cross-domain LGCMs with a Time Invariant Covariate

Exploratory modeling: Before we fit a cross-domain LGCM, determining an appropriate form of the growth curves for each domain (i.e., BMI and PBF) is necessary. Patterns of changes over time in BMI and PBF for subjects enrolled in the Fels study were modeled earlier using low-degree polynomial functions, such as linear, quadratic and cubic functions (e.g., Guo et al., 1997, 1999; Sabo et al., 2012a; Sun et al., 2012). Hence, we initially compare a series of LGCMs for each domain separately. By comparing the model fit and estimation results, we determine the form of growth curves for these two domains, and use those forms in our subsequent cross-domain analysis. For each domain, linear, quadratic and cubic growth curve models were fitted to the data, with age as the time-varying and sex as a time-invariant covariate. The models were estimated using the popular Bayesian software `JAGS` (Plummer, 2003). Model fit was evaluated using the widely applicable information criterion (or, Watanabe–Akaike information criterion,

Table 1: WAIC and LOO values corresponding to the linear, quadratic, and cubic separate-domain LGCMs models fitted to the BMI and PBF responses in the Fels data. Entries corresponding to the PBF for the cubic LGCM are absent (−), due to model non-convergence.

	BMI		PBF	
	WAIC	LOO	WAIC	LOO
Linear LGCM	17708.2	17975.5	−13772.1	**−13499.9**
Quadratic LGCM	17032.0	17602.6	**−13907.6**	−13176.5
Cubic LGCM	**16348.6**	**17315.4**	–	–

WAIC), and the leave-one-out cross-validation, LOO (Vehtari et al., 2017; Watanabe, 2010). The two fit criteria were computed using the `loo` package (Vehtari et al., 2019) in `R`.

Table 1 presents the WAIC and LOO values corresponding to the 3 model fits (linear, quadratic and cubic LGCMs), with the corresponding parameter estimates appearing in Tables A.1–A.3 in Web Supplement A, respectively. For BMI, both criteria preferred the cubic model. However, after further examining parameter estimates from the cubic LGCM, we observed that the posterior (mean) estimates for the quadratic and cubic terms were $\approx 0 (< 0.01)$ and non-significant, the variances of the two terms were also very small, and the correlations between cubic term and other random variables were not significant, too, indicating that there is minimal quantifiable inter-individual differences revealed from the quadratic and cubic terms, and they are not interrelated. Thus, it is reasonable to infer that the quadratic and cubic terms are not necessary in the model. Furthermore, we fitted individual growth curves to each individual. Overall, the linear form of growth trajectories fitted BMI scores reasonably well with the average R-squared being 0.67.

For the PBF response, we considered a LGCM under a Beta regression framework (Ferrari and Cribari-Neto, 2004; Bandyopadhyay et al., 2017), popularly used for modeling proportion responses, with a common precision parameter for all subjects. The cubic model did not converge even after 100,000 iterations. WAIC preferred the quadratic model, while LOO preferred the linear LGCM. Although WAIC is asymptotically equal to LOO, LOO is more robust in the finite sample cases (Vehtari et al., 2017). Hence, we rely on the suggestions from LOO. Similarly as above, the posterior mean estimates of the quadratic terms were ≈ 0 (< 0.01), and non-significant. Hence, we posit the linear LGCM as the most desirable for modeling PBF.

Cross-domain modeling: We thus select linear LGCMs for both BMI and PBF, and the linear form will be used in subsequent cross-domain analysis. The main structural part of the cross-domain LGCM is presented in Figure 1. Controlling for sex, we are interested in how the latent intercept I_B of BMI influences the latent intercept I_P and slope S_P for PBF, and whether and how the latent slope S_B of BMI affects the latent slope S_P of PBF.

To fit such a cross-domain LGCM to the Fels data, we encounter three major challenges. First, typical to most observational data, subjects in our analytical dataset have unique time-profiles, measured at varying time points with varying intervals. Handling individually-varying time metrics within the SEM framework is not straightforward. To circumvent this, we propose to use definition variables (Mehta and Neale, 2005), i.e., observed variables used to fix model parameters to individual specific data values, such that the factor loading matrix in SEM will be converted based on the definition variables, instead of the varying time points. Second, PBF responses are proportion data $\in (0, 1)$, where traditional normality-based linear mixed models

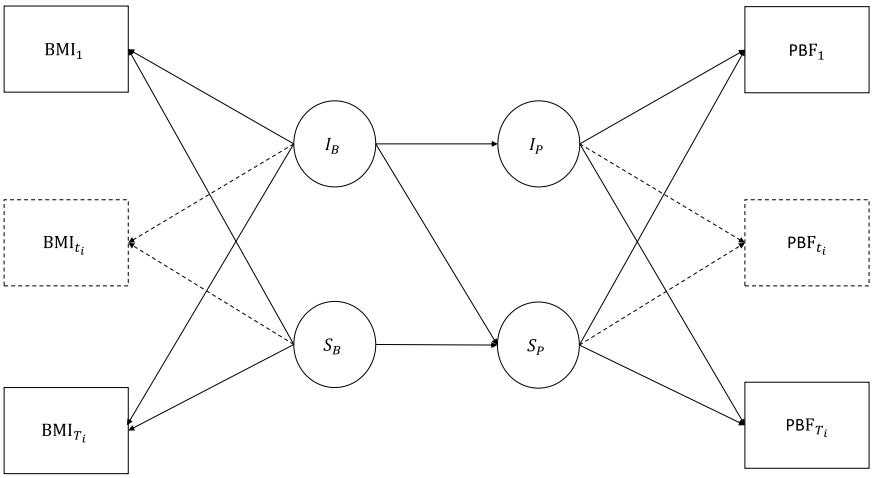

Figure 1: Path diagram of a cross-domain latent growth curve model, illustrating the association between BMI and PBF for subject i. Here, BMI_{t_i} and PBF_{t_i} denote the respective BMI and PBF measures at time t_i, with (I_B, S_B) and (I_P, S_P), the respective latent (intercept, slope) combinations.

may not provide accurate and precise parameter estimates. Hence, we employ beta regression (Bandyopadhyay et al., 2017), and introduce an added-on structure to the SEM models. Third, missing responses, a commonplace within the Fels database, have been previously handled via listwise deletion, full-information-maximum likelihood estimation (FIML), or multiple imputation (MI) methods (Guo et al., 1997, 1999; Nahhas et al., 2010; Sun et al., 2012), under missing completely at random (MCAR), or missing at random (MAR) assumptions. Both FIML and MI can provide accurate parameter estimates when data are MCAR or MAR. However, it is reasonable to assume that study subjects with higher obesity status may also have higher likelihood of attrition (being too ill to participate), leading to the missing not at random (MNAR) scenario, where there is a quantifiable relationship between the outcome variable and the propensity for missing data (Enders, 2011), even after controlling for possible correlates of missingness. Therefore, we propose to use selection models (Gomes et al., 2020) with a logit link function to model the missing data indicator within our SEM framework to obtain reasonable interpretations of the missingness. In the following, we describe our modeling framework.

Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT_i})'$ be a $T_i \times 1$ vector of observed PBF scores for individual i, where y_{it} is the observation for this individual at time t with $i = 1, \ldots, N$ and $t = 1, \ldots, T_i$. Here, N is the sample size, and T_i is the total number of measured occasions for individual i. In the BMI domain, let $\mathbf{z}_i = (z_{i1}, \ldots, z_{iT_i})'$ be a $T_i \times 1$ vector of observed BMI scores for individual i, where z_{it} is the observation for this individual at time t with $i = 1, \ldots, N$ and $t = 1, \ldots, T_i$. For modeling the proportion PBF scores $\mathbf{y}_i$, we consider a beta distribution,

$$y_{it} \sim \text{Beta}(\mu_{it}\phi_i, (1 - \mu_{it})\phi_i), \tag{1}$$

with the hyperparameters $\mu_{it}\phi_i$ and $(1 - \mu_{it})\phi_i$, respectively, such that

$$E(y_{it}) = \frac{\mu_{it}\phi_i}{\mu_{it}\phi_i + (1 - \mu_{it})\phi_i} = \mu_{it},$$

and

$$Var(y_{it}) = \frac{\mu_{it}\phi_i(1-\mu_{it})\phi_i}{(\mu_{it}\phi_i + (1-\mu_{it})\phi_i)^2(\mu_{it}\phi_i + (1-\mu_{it})\phi_i + 1)} = \frac{\mu_{it}(1-\mu_{it})}{\phi_i + 1}.$$

Here, ϕ_i can be interpreted as a precision parameter. For a fixed μ_{it}, a larger value of ϕ_i implies smaller variance of y_{it}. Since $\mu_{it} \in (0, 1)$, we conveniently use a logit link, such that

$$\text{logit}(\mu_{it}) = \mathbf{\Lambda}_{it.}\mathbf{b}_i, \tag{2}$$

where $\mathbf{\Lambda}_{it.}$ is the t-th row of the $\mathbf{\Lambda}_i$, a $T_i \times q$ factor loading matrix (q is the number of growth factors) determining the shape of the growth trajectories, and $\mathbf{b}_i$ is a $q \times 1$ vector of factor scores for individual i. The precision parameter ϕ_i can be modeled by a gamma distribution, Gamma(d_1, d_2). With the augmented structure in equations 1-2, the cross-domain LGCM can be expressed as

$$\begin{aligned} y_{it} &\sim \text{Beta}(\mu_{it}\phi_i, (1-\mu_{it})\phi_i), \\ \text{logit}(\mu_{it}) &= \mathbf{\Lambda}_{it.}\mathbf{b}_i, \\ \mathbf{b}_i &= \Theta\mathbf{c}_i + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\zeta}_i, \\ \mathbf{z}_i &= \mathbf{\Gamma}_i\mathbf{c}_i + \mathbf{u}_i, \\ \mathbf{c}_i &= \boldsymbol{\gamma} + \boldsymbol{\xi}_i, \end{aligned} \tag{3}$$

where $\boldsymbol{\beta} = (\beta_{I_P}, \beta_{S_P})'$ is the fixed effects for PBF (β_{I_P} and β_{S_P} are averages of the latent intercepts I_{Pi} and latent slopes S_{Pi}, respectively) and $\boldsymbol{\zeta}_i$ represents the random components in the random effects $\mathbf{b}_i$. Further, $\mathbf{\Gamma}_i$ is a $T_i \times l$ factor loading matrix determining the shape of the growth trajectories for the BMI scores (l is the number of growth factors), $\mathbf{c}_i$ is a $l \times 1$ vector of factor scores for individual i, $\mathbf{u}_i$ is a vector of intra-individual measurement errors, $\boldsymbol{\gamma} = (\gamma_{I_B}, \gamma_{S_B})'$ is the fixed effects for BMI (γ_{I_P} and γ_{S_P} are averages of the latent intercepts I_{Bi} and latent slopes S_{Bi}, respectively), and $\boldsymbol{\xi}_i$ is the random components in the factor score $\mathbf{c}_i$. The $q \times l$ matrix Θ quantifies the association between the two domains (i.e., PBF and BMI). Since we have selected the linear LGCMs to model the changes in PBF and BMI, $q = l = 2$. The vectors $\boldsymbol{\zeta}_i$, $\mathbf{u}_i$, and $\boldsymbol{\xi}_i$ follow multivariate normal distributions, with mean $\mathbf{0}$, and covariance matrices $\mathbf{\Sigma}_\zeta$, $\mathbf{\Sigma}_u$, and $\mathbf{\Sigma}_\xi$, respectively.

For handling individually-varying time metrics in this SEM, we resort to definition variables (Mehta and Neale, 2005). In the definition variable approach, the $\mathbf{\Lambda}_i$ and $\mathbf{\Gamma}_i$ matrices do not contain fixed values (e.g., 0, 1, 2, 3), but values based on the definition variables. For example, if there are four measurement occasions for individual i and we track a linear change pattern of $\mathbf{y}_i$ against age, the $\mathbf{\Lambda}_i$ matrix is then specified as

$$\mathbf{\Lambda}_i = \begin{pmatrix} 1 & \frac{\text{age}_{i1} - k_1}{k_2} \\ 1 & \frac{\text{age}_{i2} - k_1}{k_2} \\ 1 & \frac{\text{age}_{i3} - k_1}{k_2} \\ 1 & \frac{\text{age}_{i4} - k_1}{k_2} \end{pmatrix},$$

where, the first column defines the intercept and the second defines the linear age-based slope. The variables age_{ij}, $j = 1, \ldots, 4$, are the definition variables, and represent the age of individual

i at the measurement occasions, while k_1 and k_2 are constants to center the intercept and scale the slope, respectively. In this study, we centered age at the mean age 29.52, which yielded $k_1 = 29.52$. So, the intercept will represent predicted PBF or BMI scores at this age. The constant k_2 scales the slope according to the chosen time metric. In our data, age was measured in years. Hence, we did not need to rescale age, and thus set $k_2 = 1$. Based on this specification, the factor loading matrix $\mathbf{\Lambda}_i$ varies across individuals. The $\mathbf{\Gamma}_i$ matrix can be specified in the same way using definition variables, such that it also varies across individuals.

In practice, the cross-domain LGCM given by the set of equations (3) can be simplified by letting the precision parameter ϕ_i to be constant across individuals, such that $y_{it} \sim \text{Beta}(\mu_{it}\phi, (1-\mu_{it})\phi)$. We denote this simplified cross-domain growth curve model with a common precision parameter as Com-CLGCM, and the previous cross-domain growth curve model with individually varying precision parameters as Ind-CLGCM.

Handling missing data: Rubin (1976) distinguished three missing data mechanisms based on the process that gives rise to the missing data. They are missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR), respectively. MCAR and MAR data are called ignorable missing data because the missingness is independent of other variables, or related to only observed variables so the missingness can be ignored or explained. MNAR data are non-ignorable because the missingness is related to unobserved factors that influence the outcomes. Different missing data analytical methods should be applied accordingly to obtain reliable parameter estimates.

When data are ignorably missing, multiple imputation method (MI) can be easily applied in the Bayesian framework. Given the distribution of the missing values, Markov Chain Monte Carlo (MCMC) methods are used to iteratively impute estimates of the missing values. If missing values appear in the independent variables, we need to assume the distribution of them. When missing values only appear in the outcomes (like what we have in the Fels data), since the distribution of the outcomes are provided by the model, MI can be automatically applied.

When data are non-ignorable (e.g., our assumed MNAR specification as described earlier), joint models (Ibrahim et al., 2005) are often employed for parameter estimation to account for the unobserved factors that explain the missingness. Selection models (Kenward, 1998) are a type of joint models that focus on modeling the joint distribution of the outcome variable and the missingness indicator. Let R_{it} be the missingness indicator, such that $R_{it} = 1$ if y_{it} is missing and $R_{it} = 0$ otherwise. A possible model could be

$$\text{logit}(\Pr(y_{it} \text{ is missing})) = \text{logit}(\Pr(R_{it} = 1)) = \alpha_0 + \alpha_1 y_{i(t-1)} + \alpha_2 y_{it}. \tag{4}$$

Diggle and Kenward (1994) introduced this selection model for longitudinal analysis of non-ignorable missing data. The probability of missingness at time t depends on the outcome variables at the previous time point $(t-1)$ and the current time t. If the model is correctly specified, a significant α_1 implies MAR missingness, while a significant α_2 points to the MNAR scenario. If α_1 and α_2 are not significant, it may indicate that missingness is MCAR. Note that different forms of selection models are available in the literature. For example, the probability of missingness can be modeled as a function of the latent factors and auxiliary variables (Enders, 2011; Wu and Carroll, 1988). We use the model in Equation 4 because it can easily accommodate varying probabilities of missingness to individually-varying observation times in the Fels data.

By combining equations 3–4, we are able to model longitudinal proportion data with individually varying time metrics and non-ignorable missing values. We use Ind-CLGCM-Selection to represent the cross-domain growth curve model with individually-varying precision parameters

and an added-on selection structure. We use Com-CLGCM-Selection to represent the cross-domain growth curve model with a common precision parameter and an added-on selection structure.

In sum, for Study 1, we fit and compare the performances of the four models, (a) Ind-CLGCM, (b) Com-CLGCM, (c) Ind-CLGCM-Selection, and (d) Com-CLGCM-Selection, in the context of the Fels data. While Ind-CLGCM and Com-CLGCM can automatically handle ignorable missing data (i.e., MCAR and MAR data) within our Bayesian framework, the Ind-CLGCM-Selection and Com-CLGCM-Selection models are expected to also handle non-ignorable data (i.e., MNAR data).

With regards to prior specifications in Study 1, we used weakly informative priors for fixed effect parameters ($\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$) and parameters for the structural component Θ, such as $\beta \sim \mathrm{N}(0, 10^3)$ for $\beta \in \boldsymbol{\beta}$, $\gamma \sim \mathrm{N}(0, 10^3)$ for $\gamma \in \boldsymbol{\gamma}$, and $\theta \sim \mathrm{N}(0, 10^3)$ for $\theta \in \Theta$. Inverse Wishart priors were used for $\boldsymbol{\Sigma}_\zeta$ and $\boldsymbol{\Sigma}_\xi$, such that $\boldsymbol{\Sigma}_\zeta \sim \mathrm{IW}(2, I_2)$ and $\boldsymbol{\Sigma}_\xi \sim \mathrm{IW}(2, I_2)$, where I_2 is a 2×2 identity matrix. For the residual variance of $\mathbf{u}_i$, we typically assumed $\boldsymbol{\Sigma}_u = \sigma_u^2 I$ and used an inverse gamma prior InvGamma(0.001, 0.001) for σ_u^2. For the coefficients in the added-on selection structure, we used priors $\alpha_i \sim \mathrm{N}(0, 10^3)$ for $i = 0, 1, 2$. For the Com-CLGCM and Com-CLGCM-Selection models, we set prior $\phi \sim \mathrm{Gamma}(0.1, 0.01)$. For the Ind-CLGCM and Ind-CLGCM-Selection models, priors of ϕ_i was set as $\phi_i \sim \mathrm{Gamma}(6, 0.1)$. Bayesian software `JAGS` was used for data analysis. The length of the Markov chains was set at 60,000, with the first 30,000 iterations discarded as the burn-in. Model convergence was evaluated using Geweke's test (Geweke, 1991)

Besides WAIC and LOO, we also assessed the adequacy of model fit using posterior predictive checking (Gelman et al., 1996). Posterior predictive p (PP p) values based on a fit statistic $T(\cdot)$ were computed. Let $T(\mathbf{y}, \boldsymbol{\theta}^{(m)})$ be the test statistic computed for data $\mathbf{y}$ with parameter values at the mth iteration in the MCMC procedure. At each iteration, we generate a replicated data $\mathbf{y}^{rep}$ from the same hypothetical model based on $\boldsymbol{\theta}^{(m)}$. We then compute the fit statistic for both $\mathbf{y}$ and $\mathbf{y}^{rep}$. The PP p value can be approximated by the proportion of iterations where $T(\mathbf{y}, \boldsymbol{\theta}^{(m)}) < T(\mathbf{y}^{rep}, \boldsymbol{\theta}^{(m)})$. When the model is true, the replicated data generated from the model should be similar to the original sample. Thus, about 50% of the time a fit statistics based on the original sample will be smaller than that based on the replicated data, meaning that a perfect fit should have a PP p value of 0.5. An extreme PP p value close to 0 or 1 indicates poor fit. For SEM analysis, it is reasonable to use a cut-off value of 0.05 to reject poor fitting models (Muthén and Asparouhov, 2012). In our study, we used the mean of PBF as the fit statistic for computing PP p values.

3.2 Study 2: Cross-domain LGCMs with Additional Time-varying Covariates

Only a few studies have considered the effects of factors other than age and sex, such as blood pressure and waist circumference (e.g., Ho-Pham et al., 2015; Sabo et al., 2012b), on the longitudinal changes of BMI and PBF. Hence, in Study 2, we further control for time varying covariates (mentioned in Section 2), abbreviated as DBP, SBP, Bicep, BCimp, and waist, to quantify the relationship between BMI and PBF. After prior standardization of the time-varying covariates, we first determined the change patterns of BMI and PBF, separately, in presence of these covariates in the model. Again, after comparing linear, quadratic, and cubic growth curve models, we selected linear models for both BMI and PBF because the linear models fitted the data well, are parsimonious and are more interpretable. Similar to Study 1, model comparison using WAIC and LOO appear in Table B.1 (see, Web-supplement B), with parameter estimates cor-

responding to the linear, quadratic and cubic fits appear in Tables B.2-B.4, respectively (also, in Web-Supplement B).

Similar to Study 1, we modified the conventional cross-domain LGCM to accommodate individually varying time metrics and proportion outcomes. Our model is as follows:

$$\begin{aligned}
y_{it} &\sim \text{Beta}(\mu_{it}\phi_i, (1-\mu_{it})\phi_i),\\
\text{logit}(\mu_{it}) &= b_{0i} + b_{1i}\text{time}_{it} + \nu_1\text{BP}_{it} + \nu_2\text{Bicep}_{it} + \nu_3\text{BCimp}_{it} + \nu_4\text{waist}_{it},\\
b_{0i} &= \theta_{II}c_{0i} + \beta_{00} + \beta_{01}\text{sex}_i + \zeta_{0i},\\
b_{1i} &= \theta_{IS}c_{0i} + \theta_{SS}c_{1i} + \beta_{10} + \beta_{11}\text{sex}_i + \zeta_{1i},\\
z_{it} &= c_{0i} + c_{1i}\text{time}_{it} + \tau_1\text{BP}_{it} + \tau_2\text{Bicep}_{it} + \tau_3\text{BCimp}_{it} + \tau_4\text{waist}_{it} + u_{it},\\
c_{0i} &= \gamma_{00} + \gamma_{01}\text{sex}_i + \xi_{0i},\\
c_{1i} &= \gamma_{10} + \gamma_{11}\text{sex}_i + \xi_{1i},
\end{aligned} \tag{5}$$

where $u_{it} \sim \text{N}(0, \sigma_u^2)$, $(\zeta_{0i}, \zeta_{1i})' \sim \text{MN}(\mathbf{0}, \mathbf{\Sigma}_\zeta)$, $(\xi_{0i}, \xi_{1i})' \sim \text{MN}(\mathbf{0}, \mathbf{\Sigma}_\xi)$, $\phi_i \sim \text{Gamma}(d_1, d_2)$, and MN(.,.) denotes the multivariate normal distribution. In Equations 5, BP_{it} represents the blood pressure for individual i at time t. It is a latent variable and is indicated by diastolic blood pressure (DBP) and systolic blood pressure (SBP), as

$$\begin{aligned}
\text{DBP}_{it} &= \lambda\text{BP}_{it} + e_{1it},\ \ e_{1it} \sim N(0, \sigma_e^2),\\
\text{SBP}_{it} &= \text{BP}_{it} + e_{2it},\ \ e_{2it} \sim N(0, \sigma_e^2),\\
\text{BP}_{it} &\sim N(0, \sigma_{BP}^2).
\end{aligned}$$

DBP and SBP were not directly included in Equations 5 since they are highly correlated (cor=.67). By introducing the latent BP variable, the multicollinarity problem can be avoided. As in Study 1, we denote the above model as Ind-CLGCM. This model can be simplified by letting ϕ_i be a constant ϕ across individuals. We denote the cross-domain LGCM with a common precision parameter as Com-CLGCM.

The added-on selection structure can be incorporated into Ind-CLGCM and Com-CLGCM to model non-ignorable missing data as follows:

$$\text{logit}(\Pr(R_{it} = 1)) = \alpha_0 + \alpha_1 y_{i(t-1)} + \alpha_2 y_{it}. \tag{6}$$

By adding Equation 6 to Ind-CLGCM, we build the Ind-CLGCM-Selection model, and by adding Equation 6 to Com-CLGCM, we build the Com-CLGCM-Selection model. The performance of the four models is evaluated based on model selection criteria and parameter estimation results.

In order to conduct a Bayesian analysis, prior distributions for unknown parameters need to be specified. In Study 2, similar to Study 1, for the fixed effect parameters $\boldsymbol{\beta} = \{\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}\}$, $\boldsymbol{\gamma} = \{\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11}\}$, $\boldsymbol{\nu} = \{\nu_1, \nu_2, \nu_3, \nu_4\}$, $\boldsymbol{\tau} = \{\tau_1, \tau_2, \tau_3, \tau_4\}$ and parameters for the structural component $\boldsymbol{\theta} = \{\theta_{II}, \theta_{IS}, \theta_{SS}\}$, we used weakly informative priors, such as $\beta \sim \text{N}(0, 10^3)$ for $\beta \in \boldsymbol{\beta}$, $\gamma \sim \text{N}(0, 10^3)$ for $\gamma \in \boldsymbol{\gamma}$, $\nu \sim \text{N}(0, 10^3)$ for $\nu \in \boldsymbol{\nu}$, $\tau \sim \text{N}(0, 10^3)$ for $\tau \in \boldsymbol{\tau}$, and $\theta \sim \text{N}(0, 10^3)$ for $\theta \in \boldsymbol{\theta}$. For latent BP, prior distributions were specified as $\lambda \sim \text{N}(0, 10^2)$, $\sigma_e^2 \sim \text{InvGamma}(0.001, 0.001)$, and $\sigma_{BP}^2 \sim \text{InvGamma}(0.001, 0.001)$. Inverse Wishart priors were used for $\mathbf{\Sigma}_\zeta$ and $\mathbf{\Sigma}_\xi$, such as $\mathbf{\Sigma}_\zeta \sim \text{IW}(2, I_2)$ and $\mathbf{\Sigma}_\xi \sim \text{IW}(2, I_2)$, where I_2 is a 2×2 identity matrix. For the residual variance of z_{it}, an inverse gamma distributed prior was used, $\sigma_u^2 \sim \text{InvGamma}(0.001, 0.001)$. For the coefficients in the added-on selection structure, priors were set as $\alpha_i \sim \text{N}(0, 10^3)$ for $i = 0, 1, 2$. Additionally, we set $\phi \sim \text{Gamma}(0.1, 0.01)$ for the Com-CLGCM and Com-CLGCM-Selection models, and $\phi_i \sim \text{Gamma}(6, 0.1)$ for the Ind-CLGCM and

Table 2: Model fit for the four cross-domain linear growth curve models in Study 1.

	WAIC	LOO	PP p
Com-CLGCM	3838.4	4389.9	0.733
Ind-CLGCM	**3443.9**	**4082.6**	0.161
Com-CLGCM-Selection	3825.6	4354.9	0.739
Ind-CLGCM-Selection	**3433.2**	**4048.1**	0.163

Ind-CLGCM-Selection models. Again, the length of the Markov chains was set at 60,000, with the first 30,000 iterations discarded as the burn-in period. Model convergence was evaluated using Geweke's test. PP p values were also calculated for model assessment.

4 Application: Fels Data

4.1 Study 1

The four models (Com-CLGCM, Ind-CLGCM, Com-CLGCM-Selection and Ind-CLGCM-Selection) were fitted to the Fels data and their model fit information was summarized in Table 2. PP p values were all above 0.05, indicating that there were no concern with model fit. Both WAIC and LOO suggested that models with individually varying precision parameters (Ind-CLGCM and Ind-LCLGCM-Selection) fit data better than the models with a common precision parameter (Com-CLGCM and Com-CLGCM-Selection). We would like to note that although WAIC and LOO values for Ind-CLGCM-Selection model are the smallest, they cannot be directly compared with those values for Ind-CLGCM model. This is because when an added-on selection model is used, missing data indicator variables are created and included in the data. Therefore, models with the selection structure (Com-CLGCM-Selection and Ind-CLGCM-Selection) and models without the selection structure (Com-CLGCM and Ind-CLGCM) use different data, and thus WAIC and LOO for those two sets of models are not comparable. Based on Table 2, we may either select the Ind-CLGCM, or the Ind-CLGCM-Selection model. We conduct a simulation study in Section 5 to evaluate the robustness of Ind-CLGCM and Ind-CLGCM-Selection models against different types of missing data.

Table 3 summarizes the parameter estimation from the posterior distributions for the four models. Convergence was assessed using Geweke's test (Geweke, 1991). We observe no issues with convergences. The results from the models with a common precision parameter were slightly different from those with individually varying precision parameters, especially in the domain of PBF. Because both WAIC and LOO selected Ind-CLGCM model or Ind-CLGCM-Selection model, we focused on these two models henceforth. The results were similar from the Ind-CLGCM and Ind-CLGCM-Selection models, implying that the selection model did not add any additional information to interpret the cause of missingness. This indicated that MI probably can handle the missing data well. However, since there was a negative association between an observation and the probability of the observation to be missing ($\hat{\alpha}_2$), we may infer that the missing data mechanism underlying Fels data is MNAR, and rely more on the results from the Ind-CLGCM-Selection model. Hence, we interpret results from the Ind-CLGCM-Selection model below. The interpretation can be applied to the solution from Ind-CLGCM model as well. The parameter estimates for each domain (BMI and PBF) were similar to the results in separate growth curve modeling for the two domains, respectively. Both PBF and BMI tended to increase

Table 3: Summaries of the parameter estimates for the four cross-domain LGCMs fitted to the Fels data in Study 1. Values followed by an asterisk (*) imply the 95% credible intervals do not include 0.

		Parameter	Com-CLGCM	Ind-CLGCM	Com-CLGCM-Selection	Ind-CLGCM-Selection
Fixed effects (PBF)						
I	Intercept	β_{00}	−1.423*	−1.403*	−1.424*	−1.402*
	Sex	β_{01}	0.702*	0.683*	0.703*	0.678*
Time	Intercept	β_{10}	0.024*	0.024*	0.024*	0.024*
	Sex	β_{11}	0.002	0.001	0.002	0.001
Fixed effects (BMI)						
I	Intercept	γ_{00}	25.632*	25.619*	25.627*	25.635*
	Sex	γ_{01}	−0.534	−0.537	−0.525	−0.559
Time	Intercept	γ_{10}	0.353*	0.353*	0.352*	0.354*
	Sex	γ_{11}	−0.029	−0.029	−0.028	−0.031
Structural part						
I_{BMI} to I_{PBF}		θ_{II}	0.062*	0.061*	0.061*	0.060*
I_{BMI} to S_{PBF}		θ_{IS}	−0.002*	−0.002*	−0.002*	−0.002*
S_{BMI} to S_{PBF}		θ_{SS}	0.072*	0.070*	0.073*	0.070*
Random effects						
Var(ζ_{0i})		$\boldsymbol{\Sigma}_{\zeta}(1,1)$	0.052*	0.042*	0.053*	0.041*
Var(ζ_{1i})		$\boldsymbol{\Sigma}_{\zeta}(2,2)$	0.002*	0.002*	0.002*	0.002*
Cov(ζ_{0i}, ζ_{1i})		$\boldsymbol{\Sigma}_{\zeta}(1,2)$	0.000	0.000	0.000	0.000
Cor(ζ_{0i}, ζ_{1i})			0.019	0.005	0.021	0.005
Var(ξ_{0i})		$\boldsymbol{\Sigma}_{\xi}(1,1)$	38.339*	38.250*	38.253*	38.135*
Var(ξ_{1i})		$\boldsymbol{\Sigma}_{\xi}(2,2)$	0.085*	0.085*	0.085*	0.084*
Cov(ξ_{0i}, ξ_{1i})		$\boldsymbol{\Sigma}_{\xi}(1,2)$	1.442*	1.434*	1.438*	1.429*
Cor(ξ_{0i}, ξ_{1i})			0.798*	0.798*	0.798*	0.798*
Precision		ϕ	60.516*	–	60.504*	–
Residual		σ_u^2	1.658*	1.658*	1.660*	1.660*
Added-on selection structure						
Intercept		α_0	–	–	−2.601*	−2.612*
$y_{i(t-1)}$		α_1	–	–	0.842	0.830
y_{it}		α_2	–	–	−3.565*	−3.555*

as age increased. For PBF, female groups tended to have higher values than the male groups when age was 29.52 (at the mean age), but the slope difference between females and males were not significantly different from zero. For BMI, the effect of sex was not detected for both the initial level and the rate of change. The structural component parameter estimates indicated that (1) the initial state of BMI had positive effects on the initial state of PBF and negative effect on the rate of change for PBF; and (2) the slope of BMI had positive effects on the slope

Table 4: Model fit for the four cross-domain linear growth curve models in Study 2.

	WAIC	LOO	PP p
Com-CLGCM	−4441.5	−3964.8	0.745
Ind-CLGCM	**−4633.2**	**−4106.0**	0.093
Com-CLGCM-Selection	−4450.9	−3998.7	0.742
Ind-CLGCM-Selection	**−4644.4**	**−4124.6**	0.093

of PBF. The parameter estimates for the added-on selection structure (α_0, α_1, and α_2) imply that the probability of an observation to be missing is likely to be lower for a participant with a higher PBF score at the time of observation. The variances-covariance components suggested that both intercepts and slopes of PBF and BMI varied across subjects. The intercept and slope appeared to be highly correlated for BMI (correlation = 0.798), but the intercept and slope for PBF did not appear to be correlated.

4.2 Study 2

Similar to Study 1, the four models were fitted to the data. PP p values suggested all models had an adequate fit. Both WAIC and LOO suggested that cross-domain GCMs with individually varying precision parameters fit the data better (see Table 4). Table 5 summarizes parameter estimates from the four cross-domain growth curve models. Again, models with a common precision parameter provided slightly different results from the corresponding ones with individually varying precision parameters. However, the estimation results from models with or without the added-on selection structure are similar. We report results from the Ind-CLGCM-Selection model below. Same conclusion can be drawn based on Ind-CLGCM model.

Controlling for the time-varying covariates, PBF tended to increase as age increased for both males and females, and BMI appeared to increase only for females. Females tended to have higher PBF and BMI at age 29.52 than males when the covariates were controlled. One standard deviation increase in BP led PBF to decrease, and one standard deviation increase in Bicep, BCimp, and Waist led PBF to increase. For BMI, one standard deviation increase in BCimp led to decreasing BMI, and one standard deviation increase in BP, Bicep, and Waist led to increasing BMI. Once the time-varying covariates were controlled, the structural component parameter estimates showed that only the initial state of BMI influence the initial state of PBF. The parameter estimates for the added-on selection model (α_0, α_1, and α_2) show that the probability of an observation to be missing is more likely to be lower for a participant with a higher PBF score at the time of observation. After controlling for the time-varying covariates, the variance of intercept for BMI and PBF decreased compared to the variances in Study 1. The intercept and slope for BMI were strongly correlated (correlation = 0.55), but the correlation was not significant for PBF.

Overall, inclusion of more time-varying covariates in Study 2 led to several changes from Study 1. First, the variance of the intercept for PBF and BMI noticeably decreased after including the covariates. The variance of the intercept for BMI, in particular, decreased from 38 to 2.5 (approximately 93% decrease), indicating that the covariates explained a major portion of the variance in BMI. Second, the parameter estimates for the structural component became non-significant, except for θ_{II}. This result indicates that PBF and BMI values at age 29.52 are

Table 5: Summaries of the parameter estimates for the four cross-domain LGCMs fitted to the Fels data in Study 2. Values followed by an asterisk (*) imply the 95% credible intervals do not include 0.

		Parameter	Com-CLGCM	Ind-CLGCM	Com-CLGCM-Selection	Ind-CLGCM-Selection
Fixed effects (PBF)						
I	Intercept	β_{00}	−1.349*	−1.325*	−1.349*	−1.321*
	Sex	β_{01}	0.440*	0.414*	0.439*	0.408*
	BP	ν_1	−0.019	−0.014	−0.019	−0.014
	Bicep	ν_2	0.117*	0.114*	0.116*	0.114*
	BCimped	ν_3	0.169*	0.173*	0.169*	0.174*
	Waist	ν_4	0.343*	0.338*	0.342*	0.337*
Time	Intercept	β_{10}	0.005*	0.005*	0.005*	0.006*
	Sex	β_{11}	0.000	0.000	0.000	0.000
Fixed effects (BMI)						
I	Intercept	γ_{00}	22.724*	22.720*	22.722*	22.721*
	Sex	γ_{01}	0.978*	0.981*	0.980*	0.979*
	BP	τ_1	0.176*	0.177*	0.176*	0.177*
	Bicep	τ_2	0.915*	0.914*	0.915*	0.912*
	BCimped	τ_3	−0.827*	−0.827*	−0.828*	−0.826*
	Waist	τ_4	3.711*	3.711*	3.712*	3.714*
Time	Intercept	γ_{10}	−0.007	−0.007	−0.007	−0.008
	Sex	γ_{11}	0.016*	0.016*	0.016*	0.016*
BP	DBP	λ	1.001*	1.001*	1.001*	1.002*
	Residual	σ_e^2	0.327*	0.327*	0.327*	0.327*
	Var(BP)	σ_{BP}^2	0.673*	0.673*	0.673*	0.673*
Structural part						
I_{BMI} to I_{PBF}		θ_{II}	0.042*	0.045*	0.042*	0.045*
I_{BMI} to S_{PBF}		θ_{IS}	−0.002	−0.002	−0.002	−0.002
S_{BMI} to S_{PBF}		θ_{SS}	0.045	0.046	0.045	0.046
Random effects						
Var(ζ_{0i})		$\mathbf{\Sigma}_\zeta(1,1)$	0.040*	0.027*	0.040*	0.027*
Var(ζ_{1i})		$\mathbf{\Sigma}_\zeta(2,2)$	0.002*	0.002*	0.002*	0.002*
Cov(ζ_{0i}, ζ_{1i})		$\mathbf{\Sigma}_\zeta(1,2)$	0.000	0.000	0.000	0.000
Cor(ζ_{0i}, ζ_{1i})			0.005	−0.009	0.006	−0.010
Var(ξ_{0i})		$\mathbf{\Sigma}_\xi(1,1)$	2.475*	2.482*	2.472*	2.481*
Var(ξ_{1i})		$\mathbf{\Sigma}_\xi(2,2)$	0.007*	0.007*	0.007*	0.007*
Cov(ξ_{0i}, ξ_{1i})		$\mathbf{\Sigma}_\xi(1,2)$	0.072*	0.073*	0.072*	0.073*
Cor(ξ_{0i}, ξ_{1i})			0.554*	0.554*	0.554*	0.555*
Precision		ϕ	71.942*	–	71.925*	–
Residual		σ_u^2	0.370*	0.369*	0.370*	0.369*
Added-on selection structure						
Intercept		α_0	–	–	−2.776*	−2.777*
$y_{i(t-1)}$		α_1	–	–	0.155	0.183
y_{it}		α_2	–	–	−2.255*	−2.284

still correlated, after controlling for the four covariates, and the other relationships that were significant in Study 1 (θ_{IS} and θ_{SS}) may be due to the four covariates that were not controlled in Study 1. Finally, the effect of sex on intercept and slope became significant for BMI after inclusion of the four covariates, which is consistent with previous studies on BMI in that males and females are known to have different BMI growth patterns (Sabo et al., 2012a; Demerath et al., 2006; Guo et al., 1999).

5 Simulation Study

In this section, we conduct a simulation study to evaluate the numerical performance of the proposed cross-domain LGCMs and the robustness of the models, in light of handling ignorable and non-ignorable missing data.

The data generation mimics the Fels setup. We included one time-varying covariate, with values generated independently from N(0, 1) and one time-invariant covariate, with values drawn from a Bernoulli distribution with probability 0.5, which resembles the sex variable in the Fels data. The fixed effects for $\mathbf{y}$ were set as $\beta_{00} = -3.5$, $\beta_{01} = 1$, $\beta_{10} = 0.02$, $\beta_{11} = 0.005$, and $\nu = 0.5$. The fixed effects for $\mathbf{z}$ were set as $\gamma_{00} = 22$, $\gamma_{01} = 1$, $\gamma_{10} = 0.3$, $\gamma_{11} = -0.02$, and $\tau = 1$. For simplicity, we assumed that all individuals were measured at a common set of occasions, but the total number of measurements were allowed to vary by assuming $T_i \sim \text{Poisson}(6)$, where T_i is the number of measurement occasions for individual i. The parameters for the structural component were fixed at $(\theta_{II}, \theta_{IS}, \theta_{SS})' = (0.1, 0, 0.05)'$, while those of the individually varying precision parameter values were generated from $\phi_i \sim \text{Gamma}(8, 0.1)$, with the residual variance for BMI fixed at $\sigma_u^2 = 2$. Covariance matrices for $\mathbf{y}$ and $\mathbf{z}$ were set to be

$$\boldsymbol{\Sigma}_\zeta = \begin{pmatrix} 0.05 & 0 \\ 0 & 0.002 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_\xi = \begin{pmatrix} 3 & 0.1 \\ 0.1 & 0.01 \end{pmatrix},$$

respectively.

Sample sizes were set to $N = 400, 800$, with missing data generated under both MAR and MNAR conditions, and missingness rate set to 5% and 15%. Missing data were generated for $\mathbf{y}$ only (covariate missingness not considered). The minimum number of observations for each individual was set to be 3 and the maximum number of missing observations for each subject was set to 6 to prevent observations from being dominated by missing values and to generate simulated data similar to the Fels data. For the MAR simulation, missing values on $\mathbf{y}$ were dependent on values of $\mathbf{z}$ in that high values on $\mathbf{z}$ has the highest likelihood of missing on $\mathbf{y}$. The MNAR simulation depended the values on $\mathbf{y}$ itself. For each observation, the probability of missing on y_{it} was generated using the following equation: $\text{logit}(\Pr(R_{it} = 1|\zeta)) = \alpha_0 + \alpha_1 y_{i(t-1)} + \alpha_2 y_{it}$, with $(\alpha_0, \alpha_1, \alpha_2)' = (2, 0, -7)'$, which implies that a higher value on y_{it} is related to a lower probability of missing on y_{it}. For each condition, we simulated 500 datasets. For each dataset, we obtain parameter estimates from the Ind-CLGCM and Ind-CLGCM-Selection models via MCMC sampling, using 100,000 total iterations, with the first 50,000 discarded as burn-in. Convergence was evaluated using the Geweke's test (Geweke, 1991). Associated `R/JAGS` scripts for model implementation using simulated data are available at the `GitHub` link: https://github.com/bandyopd/GCM; see Web Supplement C.

Estimation bias and mean squared error (MSE) for each parameter were used to evaluate and compare the performance of the four models. Let ω denote a parameter, and $\hat{\omega}_l$ denote its estimate from the lth simulation replication, $l = 1, \ldots, L$. Then, the parameter estimate of ω,

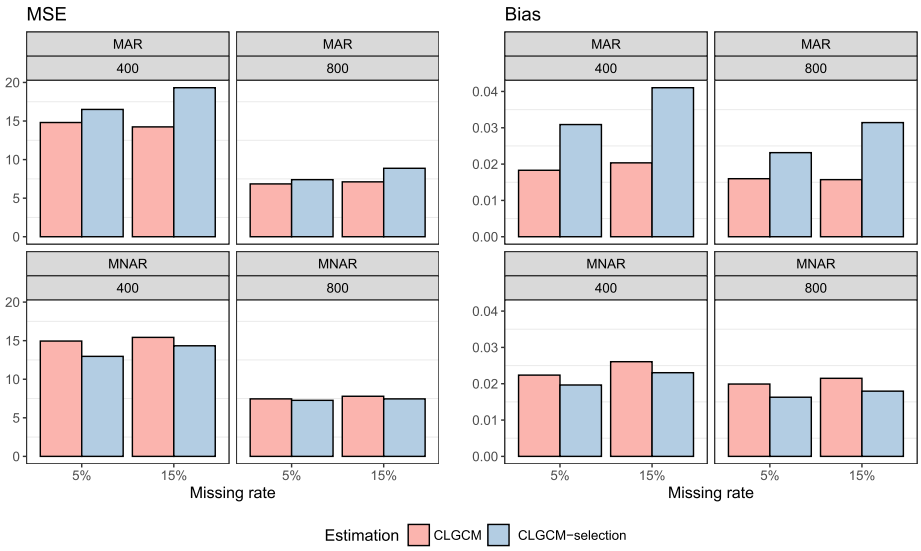

Figure 2: Averaged absolute bias and MSE for all parameters from the CLGCM and CLGCM-selection models for varying sample sizes, missing data mechanisms, and proportions of missingness.

$\hat{\omega}$, is calculated as the average of parameter estimates of L replications, given by

$$\hat{\omega} = \frac{1}{L}\sum_{l}\hat{\omega}_l.$$

The bias of $\hat{\omega}$ is $\text{bias}(\hat{\omega}) = \hat{\omega} - \omega$. The empirical standard error given as

$$\text{ESE}(\hat{\omega}) = \sqrt{\frac{1}{L-1}\sum_{l}\left(\hat{\omega}_l - \hat{\omega}\right)^2}.$$

Then, the mean squared error is calculated as $\text{MSE}(\hat{\omega}) = (\text{bias}(\hat{\omega}))^2 + (\text{ESE}(\hat{\omega}))^2$. A smaller MSE indicates a more accurate and precise estimator.

The results from this study are summarized in Figure 2, comparing the performances of Ind-CLGCM and Ind-CLGCM-Selection models, under varying conditions of missing data mechanisms (MAR and MNAR), sample sizes ($N = 400, 800$), and missing proportions (5% and 15%). We calculated the bias and MSE corresponding to each parameter ($\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, ν, τ, $\boldsymbol{\theta}$, $\boldsymbol{\Sigma}_\zeta$, $\boldsymbol{\Sigma}_\xi$), and then averaged over the parameters. Note that the original MSE values were multiplied by 10^3 for better presentation. Overall, the Bayesian estimation methods used yield very small bias and MSE. Specifically, bias values were smaller than 0.05 regardless of missing data mechanism. Bias and MSE tended to be particularly lower when missing data mechanism for data generation and model estimation were the same (i.e., MAR with CLGCM and MNAR with CLGCM-Selection in Figure 2). Bias and MSE were lower for conditions with larger sample size and lower missingness rate, but the effect of sample size appeared to be much more salient than the effect of missingness rate. In general, the small bias and MSE implies that the results in Study 1 and Study 2 are reliable. We would like to point out that the mis-specification of missing data mechanism caused larger bias and MSE, especially when missing data generation was MAR but estimated using the MNAR model (i.e., Ind-CLGCM-Selection). Because the parameter estimates from

	Multilevel Modeling	SEM	CLGCM	Com-CLGCM Ind-CLGCM	Com-CLGCM-Selection Ind-CLGCM-Selection
Proportion data				✓	✓
Individually varying time metrics	✓			✓	✓
Ignorable missingness (MCAR, MAR)	✓	✓	✓	✓	✓
Non-ignorable missingness (MNAR)					✓
Change of one variable over time	✓	✓	✓	✓	✓
Relationship between the changes of multiple variables		✓	✓	✓	✓
Latent construct with multiple indicators		✓	✓	✓	✓
Time invariant covariates	✓	✓	✓	✓	✓
Time-varying covariates	✓	✓	✓	✓	✓

Figure 3: Feature comparison between the proposed models (Com-CLGCM, Ind-CLGCM, Com-CLGCM-Selection, and Ind-CLGCM-Selection) and commonly used traditional models. Note that this is a rough guideline. Some features may be available for certain models (e.g., proportion data can be handled in the SEM framework), but commonly used software does not incorporate such features. CLGCM: cross-domain latent growth curve model.

models with and without the added-on selection structure in our studies were similar, we are not concerned much about this. In practice, if the two sets of models yield very different results, researchers need to be cautious to select the best model, and then draw conclusions.

6 Conclusions

Obesity is a complex physiological and socioeconomic issue as people did not decide to be being obesity and their weight gain is a consequence of complicated changes in the environment. BMI and PBF are two important measurements to monitor weight and assess obesity. In this article, we proposed cross-domain growth curve models in the SEM framework to investigate the longitudinal relationship between BMI and PBF and simultaneously considered the issues of proportion data, individually varying time metrics, and non-ignorable missing data. We used the definition variable approach to model the individually varying time metrics in the structural equation modeling framework, used augmented beta distribution to model proportion PBF data, and used added-on selection models to handle potential missing not at random data. Figure 3 presents a feature comparison between the proposed models, and existing traditional models.

As pointed out by the Editor, LGCMs are more readily modeled via splines (Suk et al., 2019), compared to linear/quadratic/cubic fits. However, the presence of additional covariates, and our cross-domain LGCM framework modeling both BMI and PBF, which involves proportion responses, complicates the estimation framework. Based on our extensive literature search, we were unable to find any prior work on Beta regression for cross-domain LGCM. The basis of our current choice of a simpler linear LGCM is to attain parsimony, model interpretability, and one that leads to tractable computing. We strive to consider exploring splines and other non-linear functional forms (Harring et al., 2021) within LGCMs as future work.

In light of the Fels data, we observed that BMI and PBF are associated. Comparing Study 1 with Study 2, several variations were observed in quantifying the relationship between BMI and PBF, controlling for the time varying covariates. Importantly, the average initial level and rate of change for BMI no longer predicted the average rate of change for PBF. Only the average initial levels of BMI and PBF were still associated. The effect of sex became significant for BMI after including the five covariates. In addition, a major portion of the between-persons variance in BMI can be explained by the time varying covariates. Within each study, cross-domain growth curve models without the added-on selection structure (i.e, Com-CLGCM and Ind-CLGCM) and their corresponding models with the added-on selection structure (i.e., Com-CLGCM-Selection and Ind-CLGCM-Selection) provided similar parameter estimates in our Fels data analysis. The parameters for modeling PBF appeared to be slightly influenced by the way of specifying the precision parameter in the beta distribution. As indicated by WAIC and LOO, the models with individually varying precision parameters was preferred over the models with a common precision parameter. Thus, in our simulation, we evaluated the performance of the proposed cross-domain growth curve models with individually varying precision parameters (i.e., Ind-CLGCM and Ind-CLGCM-Selection) in parameter recovery. Parameter estimates had smaller bias and MSE when the underlying missing data mechanism was correctly specified, but the two proposed models generally had very small estimation bias and MSE across all simulation conditions.

We would like to note that the performance of the selection model is sensitive to the modeling assumptions. These assumptions are usually not testable, and minor violations of the assumptions can cause biased parameter estimates (Enders, 2011). In our analysis, the added-on selection structure parameter estimates appeared to have a negative association between an observation and the probability of the observation to be missing. This result may provide an evidence that the missing data mechanism underlying Fels data is MNAR. But based on our simulation results, it may be safer to use solutions from the Ind-CLGCM model for the Fels data since both Ind-CLGCM and Ind-CLGCM-Selection models had similar performance when data were MNAR, but the Ind-CLGCM-Selection model had notably higher estimation bias when data were MAR, especially with a smaller sample size. This is inconsistent with the simulation results in Shi and Tong (2022), where a selection model was always preferred. Based on a preliminary study, we believe that the measurement reliability plays an important role in the performance of having an added-on selection structure. In our simulation study, the population parameter values were specified based on the real (Fels) data analysis, and the measurement reliability are much lower than those in Shi and Tong (2022). This may have led to a relatively worse performance of the Ind-CLGCM-Selection model. Further studies should be conducted to systematically investigate factors that may influence the performance of selection models. At this point, parameter estimates from the selection model should be interpreted with caution.

Note, although the cross-domain growth curve models in this article were specified for linear growth curves, the same idea can be applied for other nonlinear growth curves. As explained in Sterba (2014), the definition variable approach can be applied to fit polynomial, piecewise, and structured latent curves with truly individually varying time points as well. Furthermore, future studies may compare the performance of the beta regression, with beta rectangular and simplex specifications (Bandyopadhyay et al., 2017). In addition, the time invariant and time-varying covariates used in the analyses are limited to what were available in the Fels data. It is possible that other factors, such as ethnicity may also have effects on the change patterns of BMI and PBF, and the relationship between them. The Bayesian cross-domain LGCMs proposed in this paper can be applied to study different populations using other data.

Supplementary Material

Additional Tables summarizing model comparisons and parameter estimation from the two studies are available as Supplementary Materials associated with this article.

Funding

This research was supported by grants R01AG048801, R01DE031134, and R21DE031879 from the United States National Institutes of Health, and a 4-VA grant from the Commonwealth of Virginia.

References

Bandyopadhyay D, Galvis DM, Lachos VH (2017). Augmented mixed models for clustered proportion data. *Statistical Methods in Medical Research*, 26: 880–897. https://doi.org/10.1177/0962280214561093

Caussy C, Wallet F, Laville M, Disse E (2020). Obesity is Associated with Severe Forms of COVID-19. *Obesity*, 28(7): 1175. https://doi.org/10.1002/oby.22842

Demerath EW, Li J, Sun SS, Chumlea WC, Remsberg KE, Czerwinski SA, et al. (2004). Fifty-year trends in serial body mass index during adolescence in girls: The Fels Longitudinal Study. *The American Journal of Clinical Nutrition*, 80(2): 441–446. https://doi.org/10.1093/ajcn/80.2.441

Demerath EW, Schubert CM, Maynard LM, Sun SS, Chumlea WC, Pickoff A, et al. (2006). Do changes in body mass index percentile reflect changes in body composition in children? Data from the Fels Longitudinal Study. *Pediatrics*, 117(3): e487–e495. https://doi.org/10.1542/peds.2005-0572

Deurenberg P, Weststrate JA, Seidell JC (1991). Body mass index as a measure of body fatness: Age- and sex-specific prediction formulas. *British Journal of Nutrition*, 65: 105–114. https://doi.org/10.1079/BJN19910073

Diggle P, Kenward MG (1994). Informative drop-out in longitudinal data analysis. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, 43(1): 49–73.

Dulloo AG, Jacquet J, Solinas G, Montani JP, Schutz Y (2010). Body composition phenotypes in pathways to obesity and the metabolic syndrome. *International Journal of Obesity*, 34(2): S4–S17. https://doi.org/10.1038/ijo.2010.234

Duncan TE, Duncan SC (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, 35(2): 333–363. https://doi.org/10.1016/S0005-7894(04)80042-X

Enders CK (2011). Missing not at random models for latent growth curve analyses. *Psychological Methods*, 16(1): 1–16. https://doi.org/10.1037/a0022640

Ferrari S, Cribari-Neto F (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7): 799–815. https://doi.org/10.1080/0266476042000214501

Gallagher D, Visser M, Sepulveda D, Pierson RN, Harris T, Heymsfield SB (1996). How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? *American Journal of Epidemiology*, 143(3): 228–239. https://doi.org/10.1093/oxfordjournals.aje.a008733

Gelman A, Meng XL, Stern H (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6(4): 733–760.

Geweke J (1991). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: *Bayesian Statistics 4* (JM Bernardo, JO Berger, AP Dawid, AFM Smith, eds.), 169–193. Clarendon Press, Oxford.

Gomes M, Kenward MG, Grieve R, Carpenter J (2020). Estimating treatment effects under untestable assumptions with nonignorable missing data. *Statistics in Medicine*, 39(11): 1658–1674. https://doi.org/10.1002/sim.8504

Guo SS, Chumlea WC, Roche AF, Siervogel RM (1997). Age-and maturity-related changes in body composition during adolescence into adulthood: The Fels Longitudinal Study. *International Journal of Obesity*, 21(12): 1167–1175. https://doi.org/10.1038/sj.ijo.0800531

Guo SS, Zeller C, Chumlea WC, Siervogel RM (1999). Aging, body composition, and lifestyle: The Fels Longitudinal Study. *The American Journal of Clinical Nutrition*, 70(3): 405–411. https://doi.org/10.1093/ajcn/70.3.405

Harring JR, Strazzeri MM, Blozis SA (2021). Piecewise latent growth models: Beyond modeling linear-linear processes. *Behavior Research Methods*, 53: 593–608. https://doi.org/10.3758/s13428-020-01420-5

Ho-Pham LT, Lai TQ, Nguyen MTT, Nguyen TV (2015). Relationship between body mass index and percent body fat in Vietnamese: Implications for the diagnosis of obesity. *PLoS ONE*, 10(5): e0127198. https://doi.org/10.1371/journal.pone.0127198

Ibrahim JG, Chen MH, Lipsitz SR, Herring AH (2005). Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association*, 100(469): 332–346. https://doi.org/10.1198/016214504000001844

Jackson A, Stanforth P, Gagnon J (2002). The effect of sex, age and race on estimating percentage body fat from body mass index: The heritage family study. *International Journal of Obesity*, 26: 789–796. https://doi.org/10.1038/sj.ijo.0802006

Jelena J, Baltic Z, Milica G, Jelena I, Marija B, Milka P, et al. (2016). Relationship between body mass index and body fat percentage among adolescents from Serbian Republic. *Journal of Childhood Obesity*, 1(2): 10. https://doi.org/10.21767/2572-5394.100009

Kenward MG (1998). Selection models for repeated measurements with non-random dropout: An illustration of sensitivity. *Statistics in Medicine*, 17(23): 2723–2732. https://doi.org/10.1002/(SICI)1097-0258(19981215)17:23<2723::AID-SIM38>3.0.CO;2-5

Lee K, Whittaker TA (2018). Statistical power of the multiple domain latent growth model for detecting group differences. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(5): 700–714. https://doi.org/10.1080/10705511.2018.1426990

Little TD (2013). *Longitudinal Structural Equation Modeling*. Guilford press.

Makris A, Foster GD (2011). Dietary approaches to the treatment of obesity. *The Psychiatric Clinics of North America*, 34: 813–827. https://doi.org/10.1016/j.psc.2011.08.004

McQueen MA (2009). Exercise aspects of obesity treatment. *The Ochsner Journal*, 9: 140–143.

Mehta PD, Neale MC (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, 10: 259–284. https://doi.org/10.1037/1082-989X.10.3.259

Muthén B, Asparouhov T (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3): 313–335. https://doi.org/10.1037/a0026802

Nahhas RW, Choh AC, Lee M, Chumlea WMC, Duren DL, Siervogel RM, et al. (2010). Bayesian longitudinal plateau model of adult grip strength. *American Journal of Human Biology*, 22: 648–656. https://doi.org/10.1002/ajhb.21057

Nasr Eldeen SK, Al-Buni R, Al Yami A, Alali H (2017). Relationship between body mass index (BMI) and body fat percentage in a group of Saudi Arabian adults. *Annals of Public Health and Research*, 4(2): 1059.

Petridou A, Siopi A, Mougios V (2019). Exercise in the management of obesity. *Metabolism, Clinical and Experimental*, 92: 163–169. https://doi.org/10.1016/j.metabol.2018.10.009

Plummer M (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 12(125), 1–10. Vienna, Austria.

Ranasinghe C, Gamage P, Katulanda P, Andraweera N, Thilakarathne S, Tharanga P (2013). Relationship between body mass index (BMI) and body fat percentage, estimated by bioelectrical impedance, in a group of Sri Lankan adults: A cross sectional study. *BMC Public Health*, 13(1): 797. https://doi.org/10.1186/1471-2458-13-797

Roche AF (1992). *Growth, Maturation, and Body Composition: The Fels Longitudinal Study 1929-1991*. Cambridge University Press, New York.

Rubin DB (1976). Inference and Missing Data. *Biometrika*, 63(3): 581–592. https://doi.org/10.1093/biomet/63.3.581

Sabo RT, Lu Z, Daniels S, Sun SS (2012a). Serial childhood BMI and associations with adult hypertension and obesity: The Fels Longitudinal Study. *Obesity*, 20(8): 1741–1743. https://doi.org/10.1038/oby.2012.58

Sabo RT, Ren C, Sun SS (2012b). Comparing height-adjusted waist circumference indices: The Fels Longitudinal Study. *Open Journal of Endocrine and Metabolic Diseases*, 2(3): 40–48. https://doi.org/10.4236/ojemd.2012.23006

Samuels JD (2020). Obesity Phenotype is a Predictor of COVID-19 Disease Susceptibility. *Obesity*, 28(8): 1368. https://doi.org/10.1002/oby.22866

Shi D, Tong X (2022). Mitigating selection bias: A Bayesian approach to two-stage causal modeling with instrumental variables for nonnormal missing data. *Sociological Methods & Research*, 51(3): 1052–1099. https://doi.org/10.1177/0049124120914920

Sterba SK (2014). Fitting nonlinear latent growth curve models with individually varying time points. *Structural Equation Modeling*, 21: 630–647. https://doi.org/10.1080/10705511.2014.919828

Suk HW, West SG, Fine KL, Grimm KJ (2019). Nonlinear growth curve modeling using penalized spline models: A gentle introduction. *Psychological Methods*, 24(3): 269–290. https://doi.org/10.1037/met0000193

Sun SS, Deng X, Sabo R, Carrico R, Schubert CM, Wan W, et al. (2012). Secular trends in body composition for children and young adults: The Fels Longitudinal Study. *American Journal of Human Biology*, 24(4): 506–514. https://doi.org/10.1002/ajhb.22256

Vehtari A, Gabry J, Magnusson M, Yao Y, Gelman A (2019). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 2.2.0.

Vehtari A, Gelman A, Gabry J (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27: 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Watanabe S (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11: 3571–3594.

Wells JC, Fewtrell MS (2006). Measuring body composition. *Archives of Disease in Childhood*, 91(7): 612–617. https://doi.org/10.1136/adc.2005.085522

WHO (2017). World Health Organization report on prevalence of obesity among adults, BMI $\geqslant$ 30, age-standardized, estimates by country. https://apps.who.int/gho/data/view.main.CTRY2450A?lang=en. Accessed: 2023-03-31.

Wu MC, Carroll RJ (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, 44: 175–188. https://doi.org/10.2307/2531905