Multi-level modelling of longitudinal child growth data: a comparison of growth models in the Generation XXI birth cohort

Carla Moreira[1], Luís Meira-Machado[2], Maria João Fonseca[3], Ana Cristina Santos[34]

[1] CMAT, Departamento de Matemática e Aplicações, Universidade do Minho, Braga, Portugal
[2] Centre of molecular and Environmental Biology & Department of Mathematics and Applications, University of Minho, Portugal
[3] EPIUnit - Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal
[4] Departamento de Ciências da Saúde Pública e Forenses e Educação Médica, Faculdade de Medicina, Universidade do Porto, Portugal

E-mail for correspondence: carlamgmm@gmail.com

Abstract: Several methods have been used by different authors to estimate growth curves. In this paper, five growth models are compared using weight and height data from 6668 children from the Generation XXI birth cohort. The goal was to determine the model(s) that better describe the growth pattern from birth to 10 years of age using mixed-effect modelling. The study compared the fitness of four structural (Jenss-Bayley, adapted Jenss-Bayley, Berkey-Reed 1st order and Berkey-Reed 2nd order) and one non structural (cubic spline based) model. The goodness of fit of the models was examined using standard deviation of the residuals, Akaike Information Criterion and Bayesian Criterion. The adapted Jenss-Bayley and the spline based model had the better fitting for weight while for height the better models were Berkey-Reed 2nd order and the spline.

Keywords: Body height; Body weight; Generation XXI; Growth models.

1 Introduction

Weight and length/height measures are generally recorded through time by health practitioners to monitor growth and to control infants' health, especially during the first months of life. Weight and length/height growth charts have been used for decades to detect pathological growth deviations

This paper was published as a part of the proceedings of the 34th International Workshop on Statistical Modelling (IWSM), Guimarães, Portugal, 7–12 July 2019.

(WHO Multicentre Growth Reference Study Group, 2006). Over the last decades, several growth models have been developed and used in various disciplines to understand and capture general features of growth processes. The proposed modelling approaches can mainly be subdivided into non-structural and structural models. Nonstructural models do not postulate a particular form of the growth curve and are usually based on fractional polynomials and splines. These models usually fit very well the data but their parameters cannot be interpreted. Structural models imply a basic functional form and their parameters allow some biological interpretation (Singer and Willett, 2003). The growth trajectories can also be modelled using mixed multi-level models that comprise a fixed effect that represents the population growth curve and random effect that allows for individual variation around the population curve.

This study aims to compare modelling approaches using mixed effects that have previously been predominately used to model infant and early childhood growth such as Berkey-Reed models (Berkey and Reed, 1987) and those used in the late childhood period, such as the Jenss-Bayley and the adapted Jenss-Bayley models (Jenss and Bayley, 1937). The rationale behind population-based growth modelling is that, while different individuals are quantitatively different, their growth over time has a similar shape. Thus, the objective of fitting a growth curve in this instance is to quantify this common shape, but at the same time take account of the between-individual differences in growth. As well as fitting individual curves, mixed effects modelling allows for fitting of a general population curve. The fixed part of a mixed model summarizes the mean structure (general population curve) and the random component of the model allows for variations in individual growth of the children. The other advantage of using mixed effects models is that they allow for modelling of longitudinal data which have a different number of measurement occasions or where some individuals have missing outcome measurements at some points or have unequal spaced intervals between measurements occasions. This study aims to compare four structural and a non-structural model that have been shown to fit well to the infant and childhood stage in high income country settings and others (Simondon et al. 1992). The objective of this study is to find a growth model that best describes physical growth of children from birth to 10 years of age using mixed effects modelling techniques, using a large European birth cohort.

2 Methods

2.1 Study Population

The participants of this study are part of the Generation XXI birth cohort. This cohort of children was assembled between April 2005 and August 2006, after delivery, during the hospital stay, at the five public units providing obstetrical and neonatal care in the metropolitan area of Porto, Portugal. Follow-up evaluations of the cohort were carried at 4 years of age (between April 2009 and July 2011), at 7 years of age (between April 2012 and March 2014), and at 10 years of age (between July 2015 and July 2017). Of the 8647 initial cohort members, 7459 (86%), 6889 (80%) and 6392 (74%) were assessed at the 4, 7 and 10 years' follow-up evaluation, respectively. As part of the standard child care in Portugal, health professionals perform anthropometric measurements to the child and record it in a Child Health Book, in every medical visit. In all Generation XXI follow-up evaluations, mothers were asked to bring their Child Health Book in order for us to abstract data on child's weight and height measurements from birth to current age. The final sample considered for this study was 6668 children.

2.2 Growth modelling

The growth models were fitted to the data using a mixed effect modelling approach. The general structure of the models to fit is

$$y = X\beta + Zu + \epsilon \tag{1}$$

where y is the $n \times 1$ vector of the observed weith/heigth, X ia a $n \times p$ matrix of the fixed effects representing the different growth models, β is a $p \times 1$ vector of the coefficients and Z is a $n \times q$ matrix of the random effects u. The $n \times 1$ vector os errors is assumed to be multivariate normal with mean zero and variance of the matrix $\sigma_\epsilon^2 I_n$. The following general structured model was defined to test the significance of different functions of age in the random component for each growth curve indexed by k

$$y_{(k)}(t_{ij}) = f_{(k)}(t_{ij}) + h_{(k)}(t_{ij}) + \epsilon_{ij}, \tag{2}$$

where $t_{ij} \geq 0$ represents the age of child i at measurement occasion j; $y_{(k)}(t_{ij})$ represents weight or height of child i at measurement occasion j; ϵ_{ij} are the random residuals; $f_{(k)}(t_{ij})$ represents fixed effects; $h_{(k)}(t_{ij})$ represents random effects; $f_{(k)}(t_{ij})$ represents the growth curve functions such as Jenss-Bayley (Jenss and Bayley, 1937), adapted Jenss-Bayley (Botton et al, 2008), Berkey-Reed (Berkey and Reed, 1973) and splines models (Pan and Goldstein, 1998); and $h_{(k)}(t_{ij})$ is a linear function with an intercept and a slope. The random components allows for variations in the individual's child's starting measurement (intercept) and rate of growth (slope). For the weight and height models the starting measurements were child's birth weight and length/hight at birth. Individual velocities curves were computed for all children, using the estimated parameters, by derivation of the growth models. Comparisons of any nested models were done using the likelihood ratio tests based on Maximum likelihood estimation (MLE). Variance components were also estimated using MLE method. The unstructured variance-covariance option was used to estimate variance-covariance components, since it describes the variability in the individual growth of the children. Non-nested models were compared using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Residuals analysis was used to check for normality using normality plots. The statistical analysis was done using `R` using `nlme, lme4` and `SAEMIX R` packages.

3 Results

Of the 6668 participants used in the modelling weight and height, 3418 (51.3%) were males and 3250 (48.7%) were females. There were significant differences in average weight between boys and girls from 1 month to approximately 1 year. Similar trends were observed in mean height between the two sexes between birth to 3 years, with boys being on average taller than girls. There were no significant differences in mean weight or height between 3 years to approximately 8 years. At 9 years, the girls were on average heavier than boys. The graphical representation of the fitted curves with Jenss-Bayley, the adapted Jenss-Bayley, Berkey-Reed 1st and 2nd order models and a model based on a cubic spline on the observed weight and height are shown in Figure 1. The weight and height curves show some rapid weight and height gain approximately in the first year of life. Tables 1 and 2 shows that for weight and height (boys and girls), in general, the spline outperforms the other models, being the adapted Jenss-Bayley the best structural model for weight and Berkey-Reed 2nd order the best structural model for height. As expectable, the non-structural spline based model shows more flexibility when fitting weight at the early ages. Adapted Jenss-Bayley, Berkey-Reed 2nd order and spline model picks up the rapid weight gain from approximately 9 years, while the remained models are approximately linear and do not allow to represent the real weight gain.

TABLE 1. Goodness-of-fit statistics from fitting growth models for weight measurements (in Kg) from birth to 10 years of age.

	Model	res. sd	AIC	BIC	log lik	lik. ratio	p value
Boys	Jenss	0.9	251283.7	251386.3	-125630.9		
	Adapted Jenss	0.65	213049.3	213198.5	-106508.6	38244.47	$< .0001$
	Reed 1st order	1.01	266305.5	266408.1	-133141.7		
	Reed 2nd order	0.68	215726.2	215875.4	-107847.1	50589.28	$< .0001$
	Spline	0.91	214102.0	214195.0	-106641.0		
Girls	Jenss	0.98	245720.8	245822.6	-122849.4		
	Adapted Jenss	0.76	218825.9	218973.9	-109396.9	26904.95	$< .0001$
	Reed 1st order	1.07	255636.9	255738.7	-127807.5		
	Reed 2nd order	2.94	219892.9	220041.0	-109930.4	35754.05	$< .0001$
	Splines	3.26	214842.0	214933.0	-102411.0		

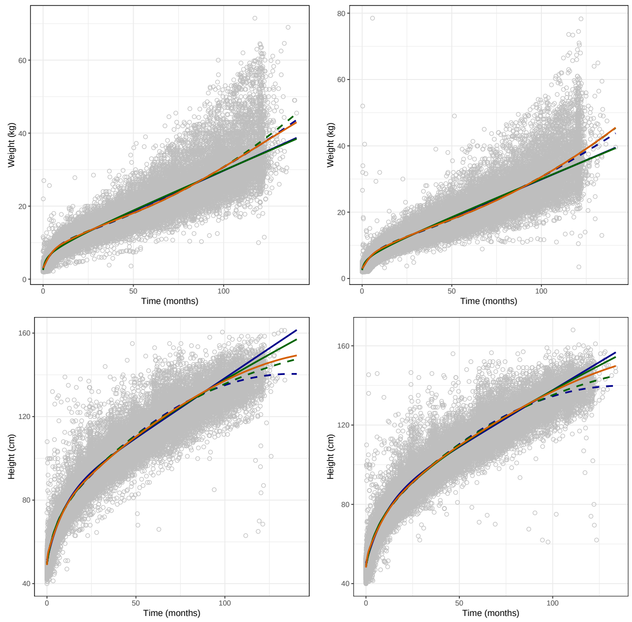

FIGURE 1. Weight and height curves for the Generation XXI birth cohort constructed using five different growth models. Top left, weight for boys and top right weight for girls. Bottom left, height for boys and bottom right height for girls. Jenss-Bayley -Solid blue line, Adapted Jenss-Bayley - dashed blue line, Berkey-Reed 1st order - solid green line, Berkey-Reed 2nd order - dashed green line and orange solid line- spline based model.

TABLE 2. Goodness-of-fit statistics from fitting growth models for height measurements (in Kg) from birth to 10 years of age.

	Model	res. sd	AIC	BIC	log lik	lik. ratio	p value
Boys	Jenss	3.02	389615.6	389716.5	-194796.8		
	Adapted Jenss	2.90	359804.6	359950.1	-179886.3	5457.048	< .0001
	Reed 1st order	3.01	385766.6	385867.6	-192872.3		
	Reed 2nd order	2.79	380409.9	380556.7	-190189.0	5366.742	< 0001
	Spline	3.13	239314.0	239407.0	-119647.0		
Girls	Jenss	3.15	365251.6	365351.7	-182614.8		
	Adapted Jenss	2.90	359804.6	359950.1	-179886.3	5457.048	< .0001
	Reed 1st order	3.21	363927.0	364027.1	-181952.5		
	Reed 2nd order	2.94	358517.8	358663.4	-179242.9	5419.199	< .0001
	Splines	3.26	214842.0	214933.0	-102411.0		

4 Conclusions

Based on AIC and BIC values and also in the standardized residuals, the best model when modelling weight and height from birth to 10 years of age has been shown to be the spline model followed by the adapted Jenss-

Bailey structured model for weight and Berkey-Reed 2nd order for height. For weight, Jenss-Bayley and Berkey-Reed 1st order showed a negative residuals suggesting overestimation of the predicted values compared to the observed measures. For height all the methods showed more variability than for weight, from approximately 60 months, being in general more accurate in the infancy. This study investigated the best models that can be used to model weight and height growth in infancy and childhood in order to ultimately study the determinants and outcomes of growth trajectories. models.

References

Berkey, C.S. and Reed, R.B. (1987). A model for describing normal and abnormal growth in early childhood. *Human Biology*, **59(6)**, 973 – 987.

Botton, J., Heude B., Maccario, J., Ducimetiere, P., Charles, M.A. (2008). Postnatal weight and height growth velocities at different ages between birth and 5 y and body composition in adolescent boys and girls. *Am J Clin Nutr*, **87**, 1760 – 1768.

Jenss, R. and Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology*, **9(4)**, 556 – 563.

Pan H. and Goldstein H. (1998). Multi-level repeated measures growth modelling using extended spline functions. *Statistics in Medicine*. **17**, 2755 – 2770.

Simondon, K.B., Simondon, F., Delpeuch, F. and Cornu, A. (1992). Comparative study of five growth models applied to weight data from congolese infants between birth and 13 months of age. *American Journal of Human Biology*, **4(3)**, 327 – 335.

Singer, J.D. and Willett, J.B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press, 2003.

WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatrica*, 95 (S450), 76 – 85