Computer Methods and Programs in Biomedicine 163 (2018) 79–85

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

New approaches to obtaining individual peak height velocity and age at peak height velocity from the SITAR model

Zhiqiang Cao[a], L.L. Hui[b,c], M.Y. Wong[a,*]

[a] Department of Mathematics, The Hong Kong University of Science & Technology, Hong Kong SAR, China
[b] Department of Paediatrics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China
[c] School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

ARTICLE INFO

Article history:
Received 16 November 2017
Revised 30 April 2018
Accepted 24 May 2018

Keywords:
Age at peak height velocity
Growth curve
Peak height velocity
Puberty
SITAR

ABSTRACT

Objective: We compared three methods for estimating the individual peak height velocity (PHV) and age at peak height velocity (APHV) from the SuperImposition by Translation and Rotation (SITAR) model.
Methods: We fitted the SITAR model using simulated data and heights of 12 girls from the Chard Growth Study and obtained individual PHVs and APHVs from three methods: the model method, the quadratic function method and the numerical method, which are available in our newly developed R package"iapvbs". The mean, interquartile range, range of biases in estimated APHV and PHV as well as the rates of warning and unreasonable cases, i.e. estimated APHVs being outside the range of age measurements, from the three methods were presented and compared.
Results: When the growth curves of all individuals were well fitted by the SITAR model, all three methods estimated individual APHVs with similarly small biases, with a few unreasonable cases (0.16%) observed when the model method was used while more computation time required for the numerical method. When the growth curves of some individuals were not very well fitted, the model method generated more unreasonable individual APHV (8.15%) and more bias in PHV and APHV, compared to those estimated by the numerical method and quadratic function method. In line with the observations from the simulated data, the real data analysis demonstrated that the numerical method generated more reliable PHV and APHV for individuals with growth curve not well fitted by the SITAR model.
Conclusion: The performance of different methods estimating individual APHV depends largely on how well the growth curves are fitted by the SITAR model. The quadratic function method is more superior when growth curves of all individuals are well fitted by the SITAR model; otherwise, the numerical method should be adopted for getting most robust estimates of PHV and APHV. The model method generates unreasonable APHV estimates, particularly when the growth curves are not well fitted.

1. Introduction

Puberty is a process of biological maturation, involving a growth spurt and development of secondary sexual characteristics. Individuals begin puberty at different times and progress through puberty at different rates. Pubertal growth and development can be described in terms of timing (i.e. age at which an individual reaches a certain pubertal stage), velocity (i.e. how fast an individual progresses through puberty) and duration (i.e. the period of time between onset and completion or between any two stages of puberty). These pubertal parameters vary greatly across individuals [1]. Studying their determinants and long-term health consequences can provide useful insights into disease aetiology [2] and prevention [3].

Peak height velocity (PHV) and age at peak height velocity (APHV) are useful and objective markers for respectively the velocity and timing of puberty for both boys and girls and can be assessed mathematically. Pubertal growth in stature typically starts with an acceleration phase, followed by a deceleration phase, and finally a cessation of growth with the closure of epiphyses in long bones [4]. Such a growth velocity curve can be characterized using different mathematical models (Jenss and Bayley [5]; Berkey and Reed [6]; Verbeke and Molenberghs [7]; Beath [8]). Recently, Cole et al. [9] derived a growth curve analysis called SuperImposition by Translation and Rotation (SITAR) based on the shape invariant model [8]. This new analysis involves fitting the following model

* Corresponding author.
E-mail address: mamywong@ust.hk (M.Y. Wong).

https://doi.org/10.1016/j.cmpb.2018.05.030

with fixed and random effects to a set of height growth curves.

$$y_{ij} = \alpha_0 + \alpha_i + h\left(\frac{x_{ij} - m_x - \beta_0 - \beta_i}{\exp(-\gamma_0 - \gamma_i)}\right) \quad (1)$$

where x_{ij} is the ith individual at the jth age measurement; y_{ij} is the corresponding height measurement; m_x is the overall mean of x_{ij}; $\alpha_0, \beta_0, \gamma_0$ are fixed effects; $\alpha_i, \beta_i, \gamma_i$ are random effects for the ith individual; and $h(z)$ is the natural cubic spline curve [10] of the growth variable regressed on z (the transformation of age). The SITAR model summarizes the growth curves of a group of individuals (or a population) in terms of an average curve plus a set of three random effects for each individual that define how his/her growth curve differs from the average. Specifically, the three subject-specific random effects describe the size (α, shift of the growth curve), tempo (β, relative timing of peak velocity) and velocity (γ, relative growth rate) of each individual's growth.

The R package "sitar" [11] can fit model (1) and compute population PHV and APHV. However, it does not give individual PHV and APHV directly. Based on the assumption that the model fits the growth curves of all individuals closely, the individual APHV can be calculated from the population APHV (obtained using the "plot.sitar" function in the R package "sitar") and parameter estimates from the SITAR model (we refer to this method as the model method). In this paper, we propose two other methods, the quadratic function method and the numerical method, to obtain individual PHVs and APHVs based on the SITAR model and develop an R package "iapvbs"[1] for the related calculations. We also compare and contrast the performance of these three methods in calculating individual PHVs and APHVs through simulations and with longitudinal height data from 12 girls in the Chard Growth Study [11].

2. Materials and methods

2.1. Data description

2.1.1. Simulated data

To compare these three methods, we simulated growth data of 100 individuals using Model 2 from the Preece and Baines family of growth functions [12]. The function of the model and the related parameter settings, which are similar to those for the data generating process in Simpkin et al. [13], are described in Appendix A.1. Age measurements are sampled in regular and irregular occasions. For the regular measurement occasion, each individual has one age measurement per year for ten years from 10.00 to 19.99 years in a fully balanced design; that is, the jth age measurement of the ith individual is set as $Age_{ij} = 10 + 0.01^*(i-1) + (j-1)$. For the irregular measurement occasion, age measurements were randomly selected from the Uniform distribution between 10.00 and 19.99 years with different intervals between any two age measurements of an individual and different age intervals between any two individuals. Then, age measurements of an individual were deleted randomly resulting in individuals having anywhere between 5 and 10 age measurements each. For each age setting, two levels of error variance in the measurement of height ($\sigma_e^2 = 0.25$ cm, $\sigma_e^2 = 1$ cm) are considered. We conducted 1000 simulations under four scenarios: the regular measurement occasion with $\sigma_e^2 = 0.25$ cm, the regular measurement occasion with $\sigma_e^2 = 1$ cm, the irregular measurement occasion with $\sigma_e^2 = 0.25$ cm and the irregular measurement occasion with $\sigma_e^2 = 1$ cm.

2.1.2. Analysis of the heights dataset

We took advantage of the longitudinal dataset "heights" in R package "sitar", which includes 124 measurements of 12 girls between 8 and 16 years old who had their heights measured twice per year. According to the description of this small dataset, the 12 girls participated in the Chard Growth Study and the cohort was studied in Buchanan et al. [14].

2.2. The model method with the "sitar" package

2.2.1. Calculation of APHV from the model method

In essence, the R package "sitar" uses the cubic spline to fit model (1). For a piece of the spline covering the APHV of the ith individual, model (1) is equivalent to

$$y_{ij} = \alpha_0 + \alpha_i + \delta_0 + \delta_1 z + \delta_2 z^2 + \delta_3 z^3 \quad (2)$$

where $z = \frac{x_{ij} - m_x - \beta_0 - \beta_i}{\exp(-\gamma_0 - \gamma_i)}$ and x_{ij} is a point on this piece. By the definition of APHV, equating the second derivative of Eq. (2) with respect to x_{ij} to zero will give the APHV of the ith individual (denoted by apv_i). The population APHV (denoted by apv_0) can be achieved similarly by ignoring random effects and treating x_{ij} as the population age. After simplification, we can build the relationship between apv_i and apv_0; that is,

$$apv_i = \frac{apv_0 - m_x - \beta_0}{\exp(\gamma_i)} + m_x + \beta_0 + \beta_i \quad (3)$$

Using apv_0 (obtained from the "plot.sitar" function) and $\beta_0, \beta_i, \gamma_i$ (obtained from the "sitar" function in the R package "sitar"), the APHV of the ith individual can be obtained from Eq. (3). We call this method the model method and define an individual's APHV from Eq. (3) as apv_i^{model}. The velocity corresponding to apv_i^{model}, denoted by pv_i^{model}, is predicted using the "predict.sitar" function in the R package "sitar".

2.2.2. The problems

According to Eq. (3), apv_i depends on fixed effect parameter β_0 and random effect parameters β_i,γ_i. By the theory of the SITAR model, the growth curve of each individual can be obtained from the population growth curve through three random effect parameters. The poor fit of an individual's growth curve by the SITAR model indicates that the estimates of β_i and γ_i are not accurate. Thus, the accuracy of apv_i computed from Eq. (3) is questionable.

2.3. The r package "iapvbs"

2.3.1. Extrapolating heights and growth velocities

Estimating an individual's PHV and APHV requires subject-specific parameters from a well-fitted SITAR model, and estimations of these parameters can be obtained from the "sitar" function. The SITAR model with the smallest deviance, residual standard deviation and BIC (Bayesian information criterion) is chosen as the best one. After the best SITAR model is fitted, apv_i^{model} and pv_i^{model} can be obtained easily. As reference, this method is also available in our newly developed R package "iapvbs".

To estimate an individual's PHV and APHV accurately, frequent height measurements during the pubertal period are required. The "exdata" function in the R package "iapvbs" produces a data frame with interpolated age measurements and a subject identifier. After the best SITAR model is fitted, the corresponding heights and growth velocities of interpolated age measurements are then obtained using the "predict.sitar" function. Based on the interpolated ages and predicted velocities, the quadratic function method is derived. And, based on the interpolated ages and predicted heights, we derive the numerical method.

The performance of both the quadratic function method and the numerical method depends on the number of interpolated age measurements. The "getapv" function from the "iapvbs" package

[1] Available at https://github.com/Zhiqiangcao/iapvbs