NCCOR WORKSHOP

Revisiting Rapid Infant Growth: Assessing Growth in Children Under Age 2

Presentation: Research in Rapid Infant Growth and Early Life Obesity Risk- Statistical Perspective

NCCOR

Presentation: Researchers in Rapid Infant Growth and Early Life Obesity Risk- Statistical Perspective

Ahmed Elhakeem, MPH, PhD
University of Bristol

Moderator
Charles Wood, MD, MPH, FAAP,
Duke University School of Medicine

NCCOR WORKSHOP

Revisiting Rapid Infant Growth: Assessing Growth in Children Under Age 2

Modelling Infant Growth Trajectories

Ahmed Elhakeem, PhD
Bristol Medical School, University of Bristol

NCCOR

Outline

- Overview with examples of some approaches to model growth:
  - Linear mixed effects (LME) models

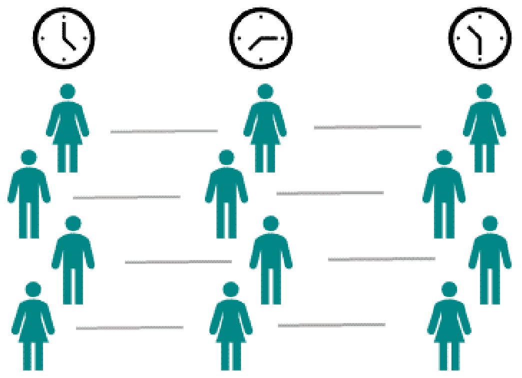
    LME - linear splines

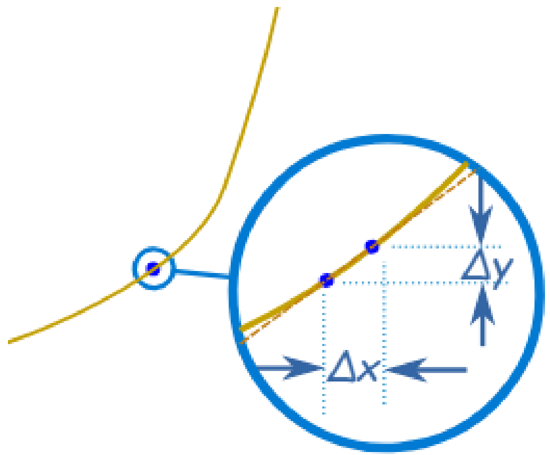
    LME - natural splines

    LME - P-splines
  - SITAR
  - Functional data analysis (FPCA)

Growth modelling – analysis of repeated measures

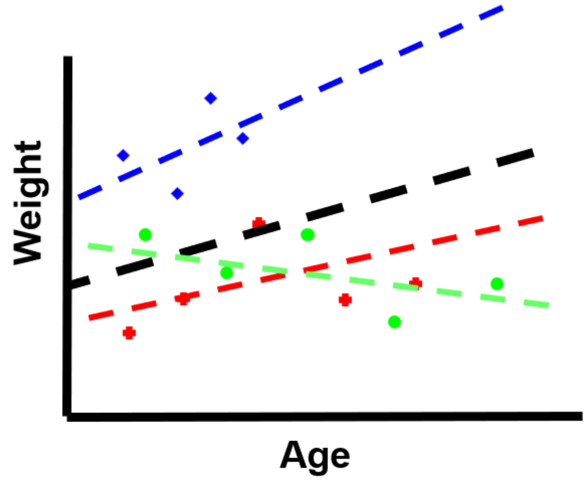
- Helps us understand how people grow and develop over time
- Identify adverse changes in growth and when these might occur
- Can improve understanding of the factors influencing trajectories and effects of adverse trajectories on later health outcomes
- Can be used to estimate growth velocity, and other curve dynamics (e.g., by calculating derivatives from the fitted distance curves)

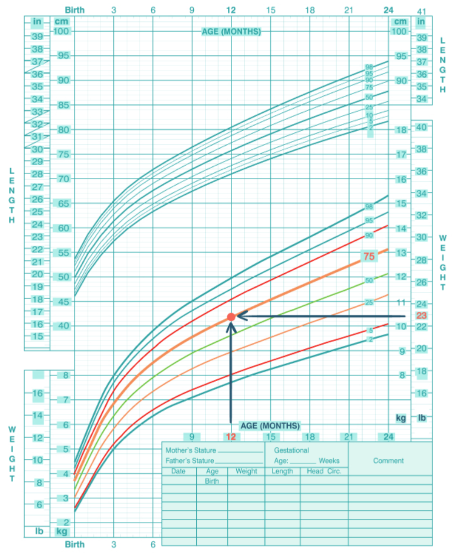




Linear mixed effects model (LME)

- Repeated outcome is evaluated by a linear combination of fixed effects and random effects.
- Can estimate a population-average trajectory (line) as 'fixed effects' as well as variation of individual trajectories around this average as 'random effects.'
- Assumes *linear change* in the outcome (e.g., weight) with increasing time (e.g., age).
- In many cases, change is not linear (e.g., infant growth – can allow for it by including linear combinations of nonlinear terms for age).

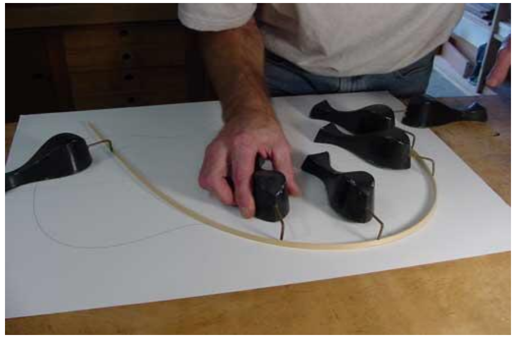


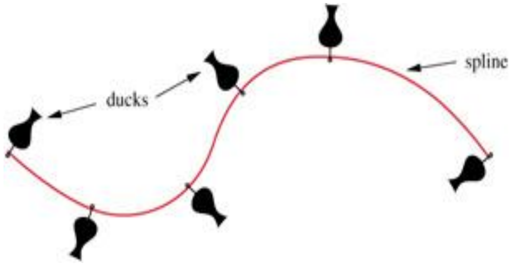


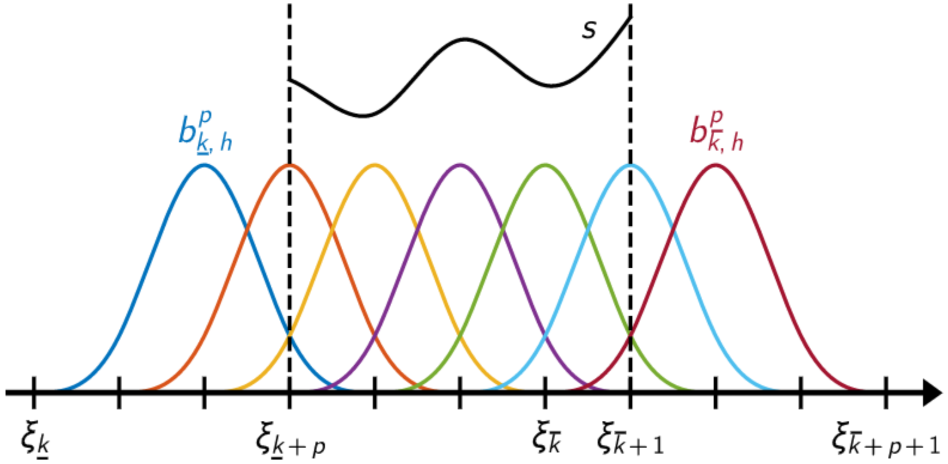
Source: WHO Child Growth Standards

Splines

- Physical crafting tools (e.g., for making boats) / statistical functions to reproduce flexible shapes.
- A spline function is a set of smoothly joined polynomial pieces. Each segment:
  - covers a specific area of the x-axis (local control)
  - is scaled by its own coefficients and
  - is defined by knots – place where segments meet
- Number and location of knots must be specified by the user.





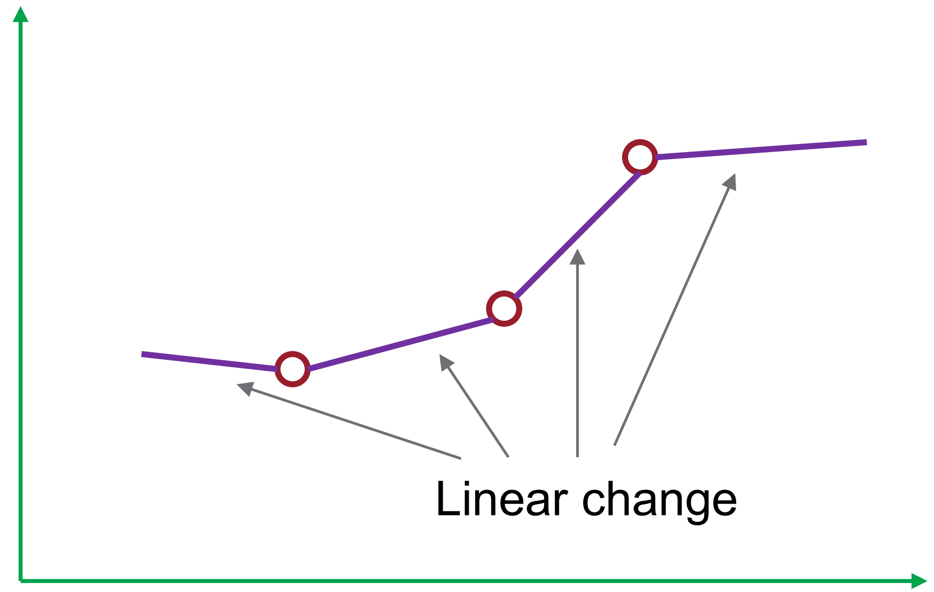


Uniform cubic B-splines forming a cubic spline.

https://bsplines.org/

Linear spline

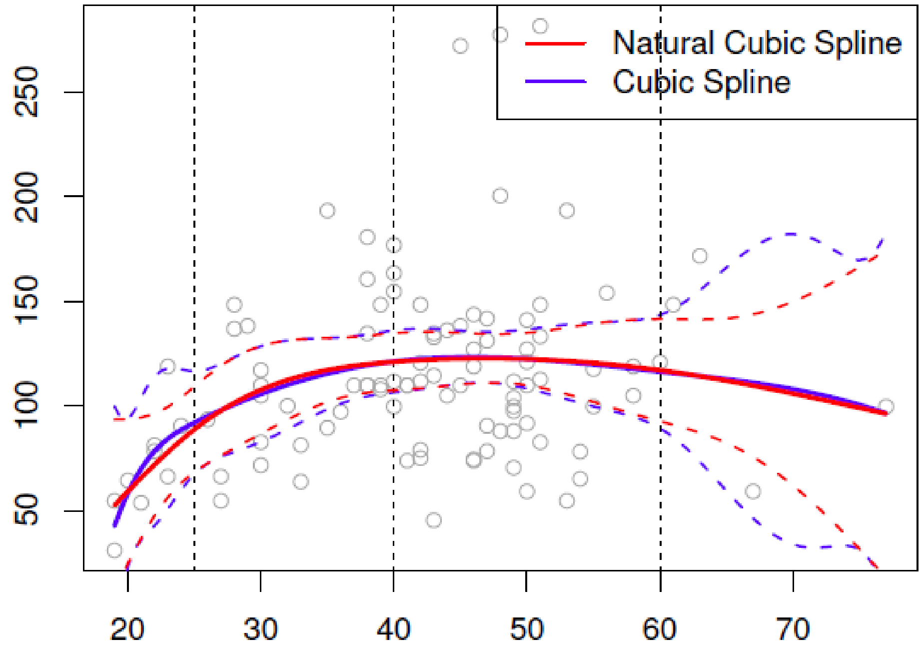
- Connected straight lines with a change in slope at each knot
- Divides a predictor (e.g., age) into mutually exclusive intervals joined at knot points
- Linear growth rate is described over age windows
- Can have unrealistic sharp changes in the trajectory



Howe LD et al. Linear spline multilevel models for summarising childhood growth trajectories: A guide to their application using examples from five birth cohorts. *Stat Methods Med Res*. 2016

Natural / restricted cubic spline

- Cubic spline: set of connected cubic polynomials with slope and continuity constraints at each knot
- Continuous 1st (velocity) and 2nd (acceleration) derivatives – can differentiate the growth curve to get a velocity curve, and identify the nature of the peak
- A natural spline is a cubic spline with additional constraints of linearity before (usually) the first knot and after the last knot; allows trajectory to be less erratic at the tails.
- Natural spline models growth by cubic polynomials between internal knots, and with a line before the first knot and after the last knot.

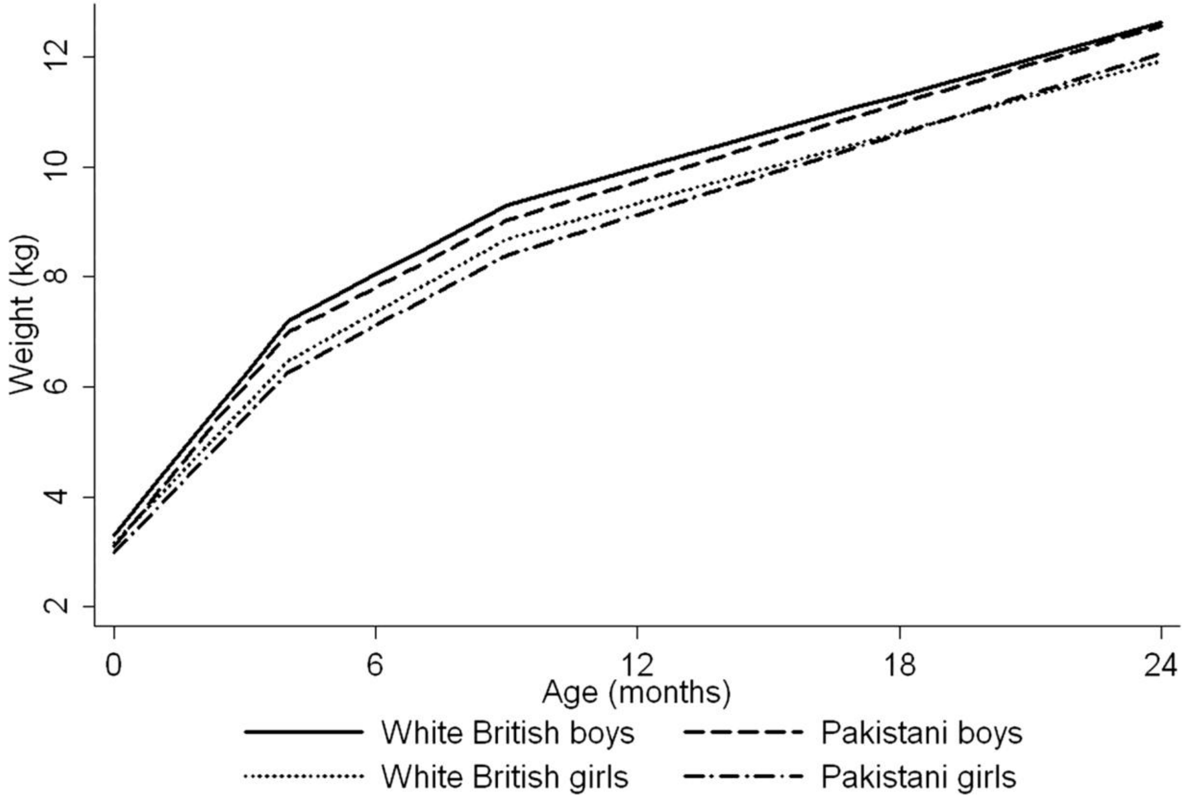


Elhakeem A. et al. Using linear and natural cubic splines, SITAR, and latent trajectory models to characterise nonlinear longitudinal growth trajectories in cohort studies. *BMC Med Res Methodol*. 2022

Choosing number and location of knots

- Flexibility/smoothness of splines is determined by the number and position of knots.
- For natural splines, knot position is less important than number (both important for linear splines).
- Common approaches to knot placement:
  - Quantile or equidistant placement
  - Placing the knots at the mean age of data collection / where data are available
  - Subject knowledge: e.G., To test sensitive periods
- Model comparison and selection can be done by a combination of informal checking and formally testing difference between fitted model.

Linear spline LME model example



median of 7 weight measures per child

Table: Mean difference in weight growth rates (kg/mo) in each time period for Pakistani vs. White British kids

Time period	Mean (SD) unadjusted growth rate for white British group	Unadjusted difference between Pakistani and white British groups		Adjusted for gestational age	
		Mean difference	95% CI	Mean difference	95% CI
Boys					
Birth	3.30 (0.57)	−0.21	(−0.29 to −0.12)	−0.24	(−0.31 to −0.16)
0–4 months	0.98 (0.17)	−0.00	(−0.03 to 0.03)	−0.00	(−0.03 to 0.03)
4–9 months	0.42 (0.12)	−0.01	(−0.04 to 0.01)	−0.01	(−0.04 to 0.02)
9–24 months	0.22 (0.04)	0.01	(0.00 to 0.03)	0.01	(0.00 to 0.03)
Girls					
Birth	3.16 (0.52)	−0.18	(−0.26 to −0.10)	−0.18	(−0.25 to −0.11)
0–4 months	0.83 (0.17)	−0.01	(−0.04 to 0.02)	−0.01	(−0.04 to 0.02)
4–9 months	0.44 (0.12)	−0.02	(−0.04 to 0.01)	−0.02	(−0.04 to 0.01)
9–24 months	0.22 (0.04)	0.03	(0.02 to 0.04)	0.03	(0.02 to 0.04)

NCCOR

Fairley L et al. Describing differences in weight and length growth trajectories between white and Pakistani infants in the UK: analysis of the Born in Bradford birth cohort study using multilevel linear spline models. *Arch Dis Child*. 2013

Natural spline model examples

- 321 infants were enrolled in South Asian (SA) sites
- 6 h of birth and at 1, 2, 3, 4, 5, 12 and 30 days of age
- *Quantile* regression with natural spline

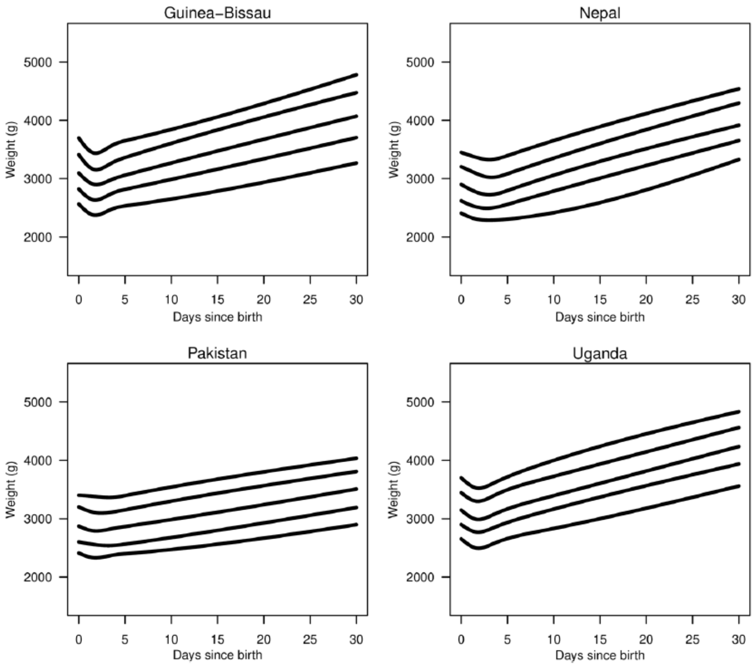
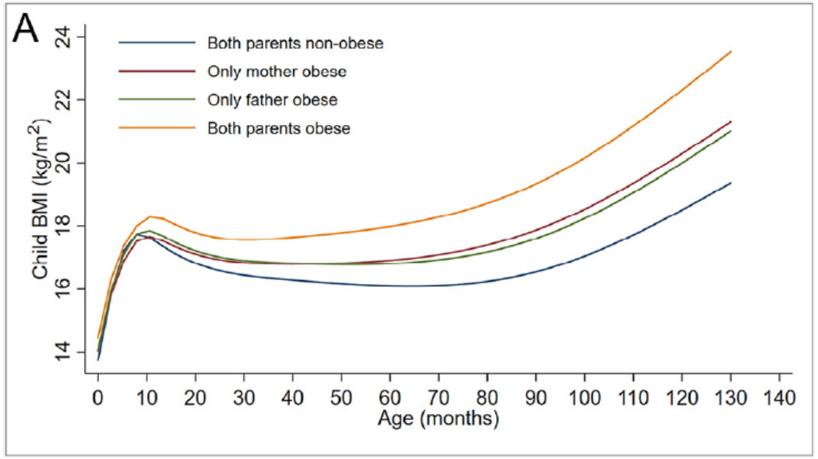


FIGURE 1 Infant weight over the first 30 days by country of birth depicted as the 10th, 25th, 50th (median), 75th and 90th percentiles of weight at each time point

Flaherman VJ et al. *Matern Child Nutr*. 2022

- Project Viva
- LME with natural spline
- Predicted values -> derivative (velocity) -> bmi peak/rebound

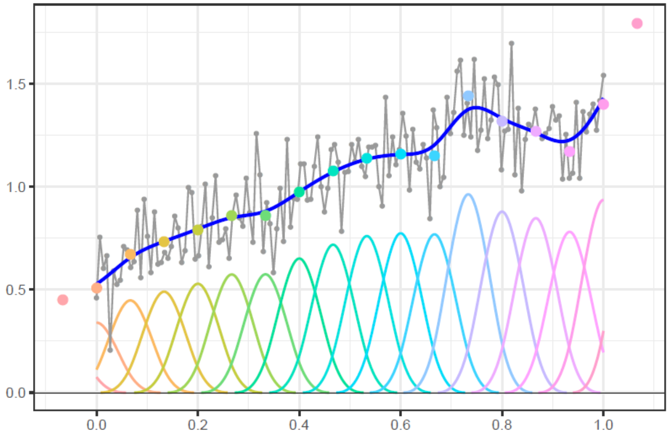


Aris IM et al. J Pediatr. 2018

	Crude analyses		
		95% CI	
Characteristics	B (kg/m^2)	Low	High
Magnitude at peak			
Parental obesity			
Both parents non-obese	ref	=	=
Only mother obese	0.2	–0.04	0.5
Only father obese	0.3	0.03	0.5
Both parents obese	0.8	0.4	1.1

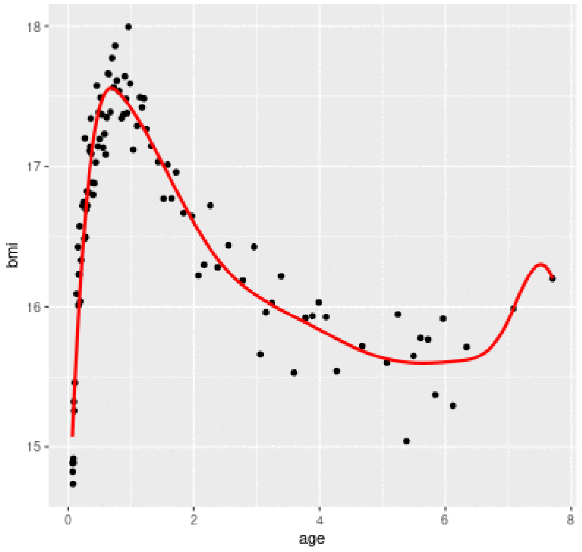
P-splines

- Splines with a penalty applied on their basis coefficients
- Do not need to specify the number of knots - decide roughly how large a basis dimension is fairly certain to provide just adequate flexibility and use that
- A smoothing parameter (lambda) penalizes the basis coefficients to avoid overfitting – can be estimated as part of the model (e.g., via REML)
- Have been used for growth chart development but not many applications in longitudinal growth curve analysis – software limitations / perceived complexity



The joys of P-splines - book

The core idea of P-splines: a sum of B-spline basis functions, with gradually changing heights.

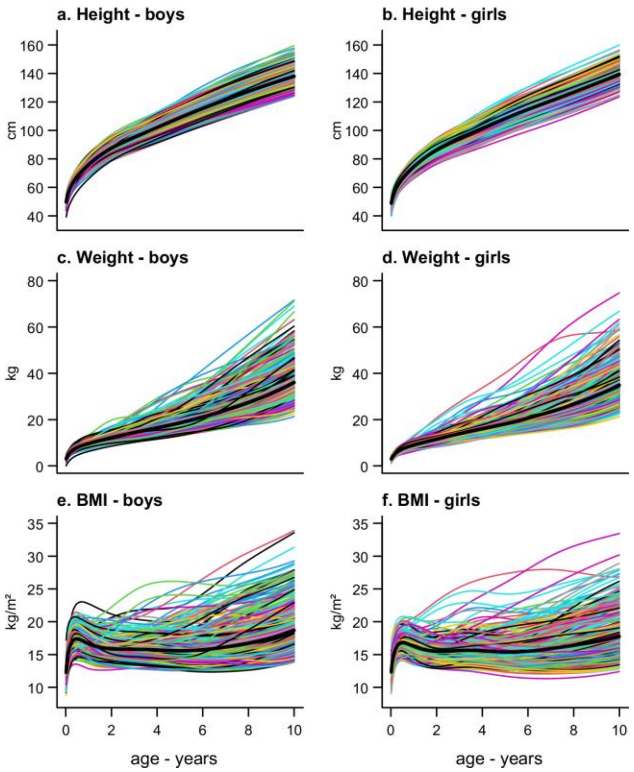


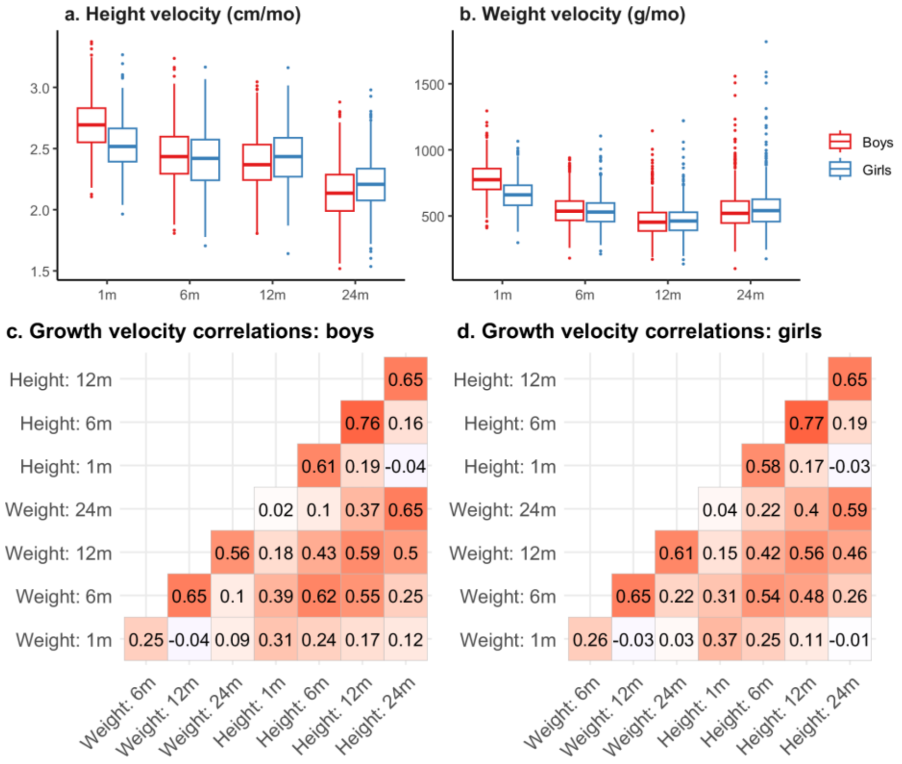
P-spline mixed effects model guide and R package (psme)

GUSTO
GROWING UP IN SINGAPORE TOWARDS HEALTHY OUTCOMES

https://github.com/LongitudinalModelling/psplines

https://cran.r-project.org/web/packages/psme/





Hernandez MA et al. Capturing infant and child growth dynamics with P-splines mixed effects models. *Int J Obes* 2026

NCCOR

Birth weight percentile and infant growth velocity

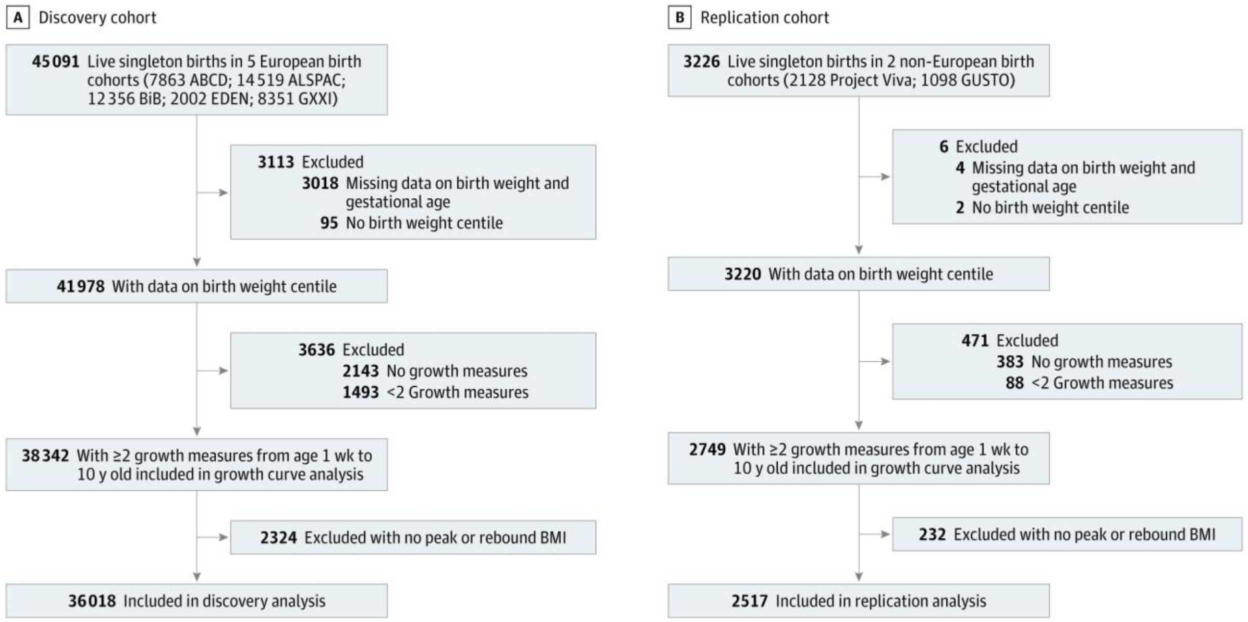
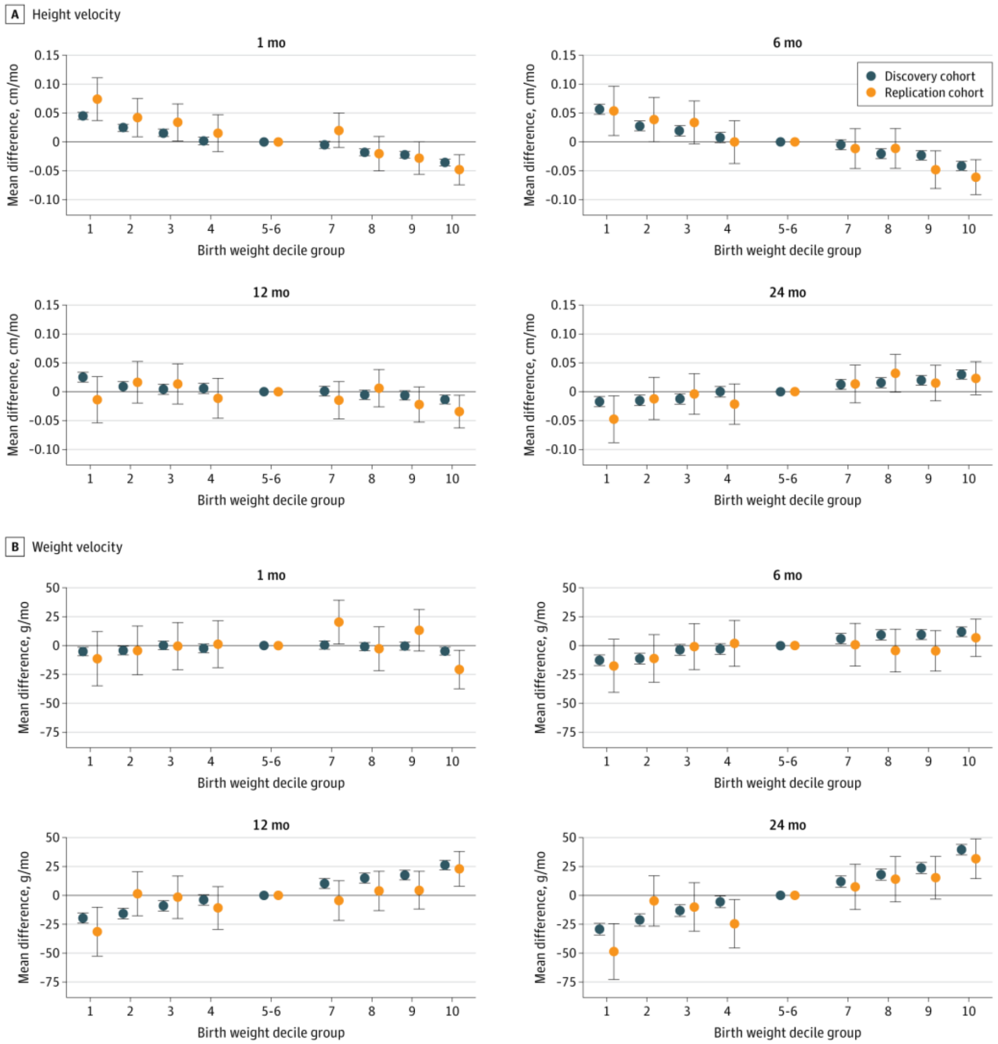


Fig. Mean difference in infant height and weight velocity at age 1, 6, 12, and 24 months by birth weight decile group



Hernandez MA et al. Birth Weight Percentiles and Infant and Child Growth Dynamics. *JAMA Netw Open*. 2026;9(7):e2622218.

Birth weight percentile and infant growth velocity

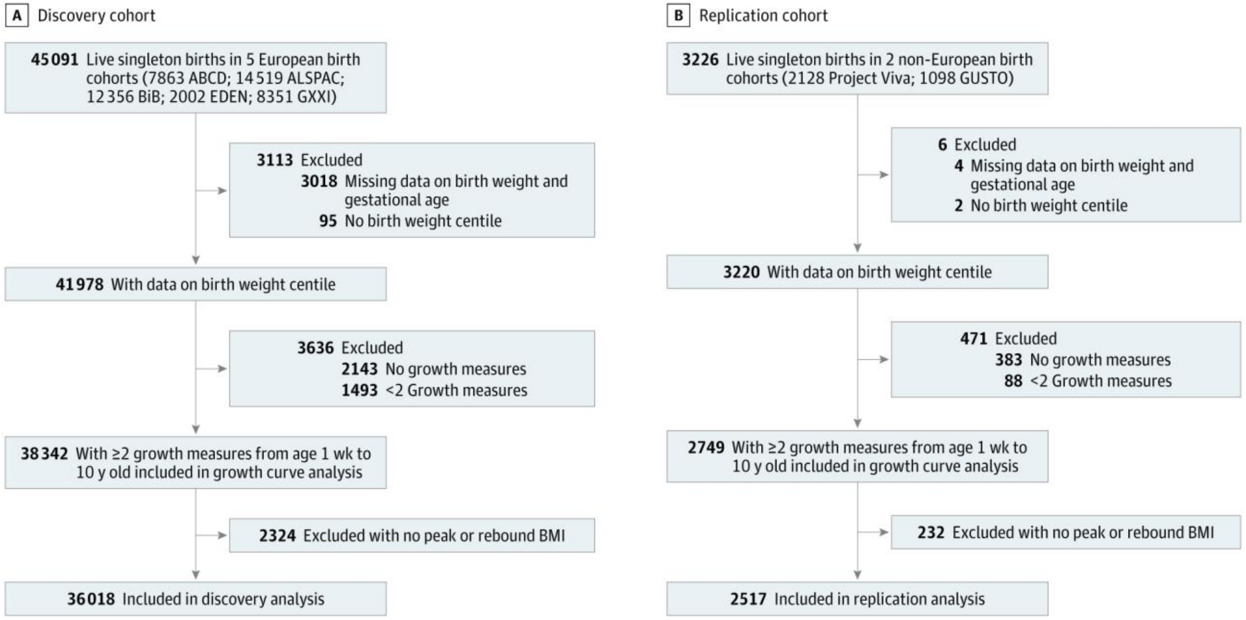
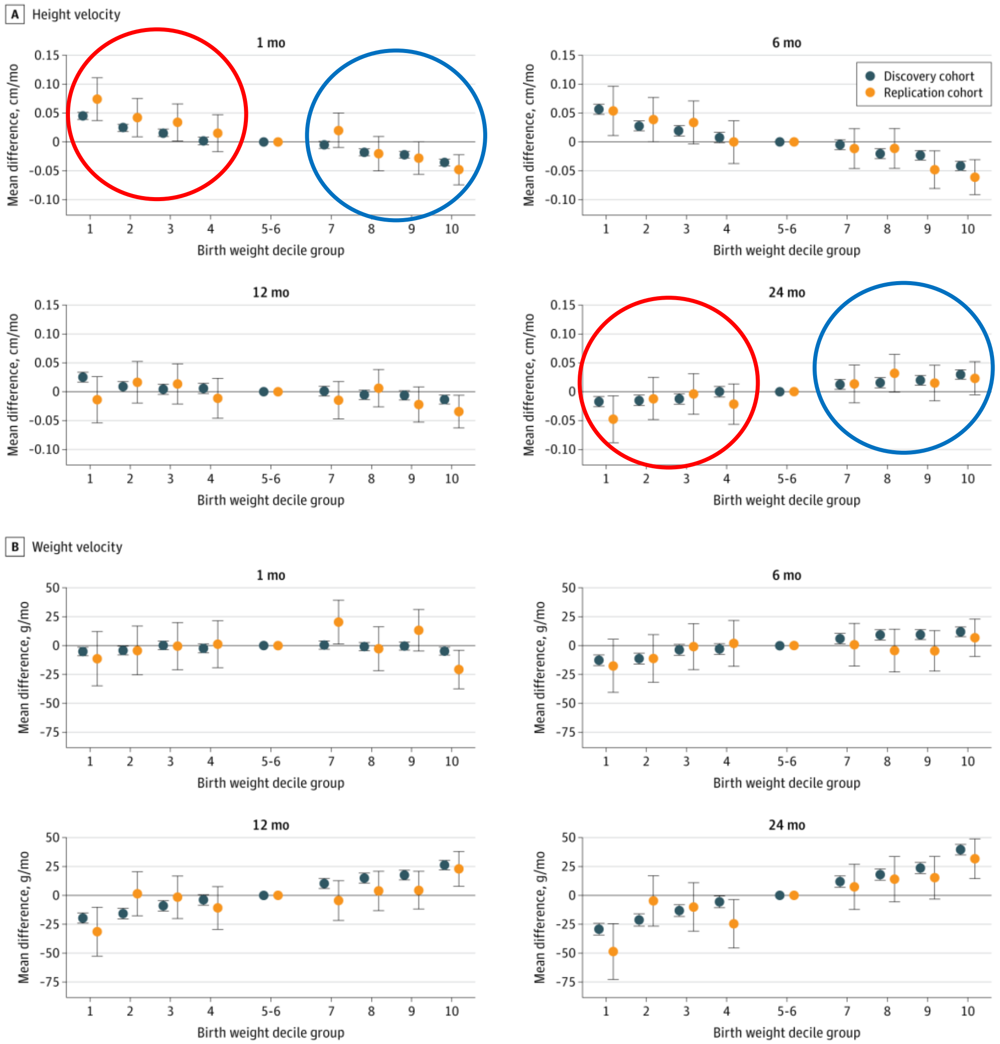


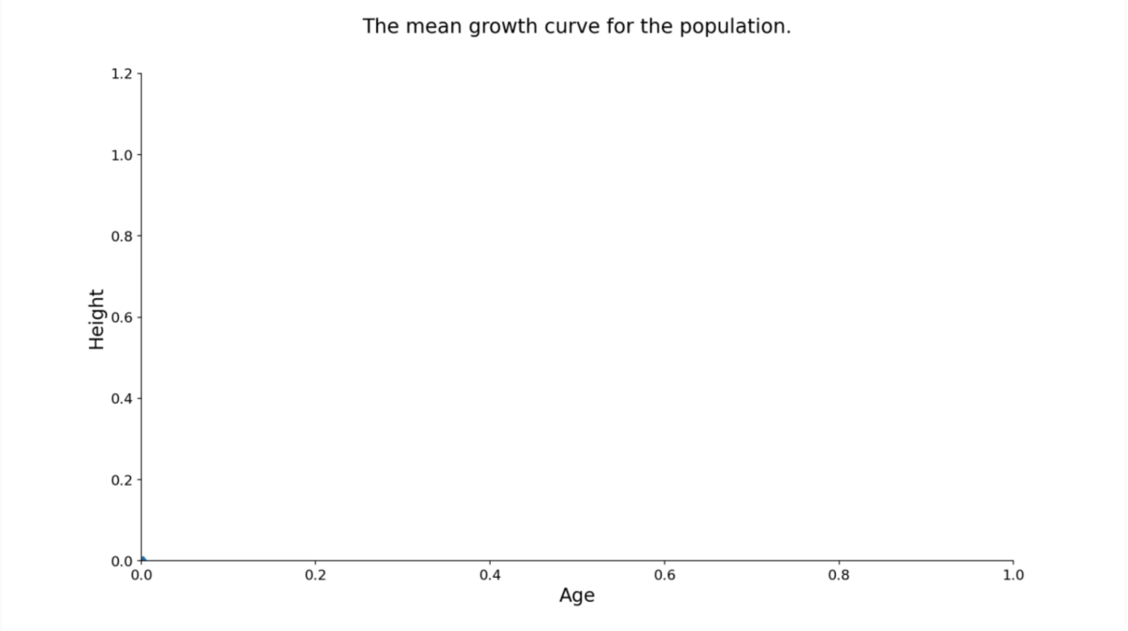
Fig. Mean difference in infant height and weight velocity at age 1, 6, 12, and 24 months by birth weight decile group



Hernandez MA et al. Birth Weight Percentiles and Infant and Child Growth Dynamics. *JAMA Netw Open.* 2026;9(7):e2622218.

SITAR

- An Indian musical instrument / a shape-invariant growth model
- Nonlinear in the coefficients
- Fits one (natural spline) curve (mean / population)
- Mean curve is shifted and scaled by 3 random effects (*a-size, b-timing and c-intensity*) to fit any individual curve
- Individual curves are all versions of the mean curve



Cole TJ et al. SITAR--a useful instrument for growth curve analysis. *Int J Epidemiol.* 201

Using SITAR to estimate peak/rebound BMI and infant peak weight velocity: analysis of 2 birth cohorts

- Weight from 0–6 months
- Weight ~ log(age+0.01)
- Compare SITAR models with 2–6 df
- Estimate infant PWV and APWV (sitar::get_peak()):
- Predicted values -> smoothed first derivatives -> turning point (sign change) -> quadratic refinement -> PWV

GUSTO
GROWING UP IN SINGAPORE TOWARDS HEALTHY OUTCOMES

PROJECT Viva

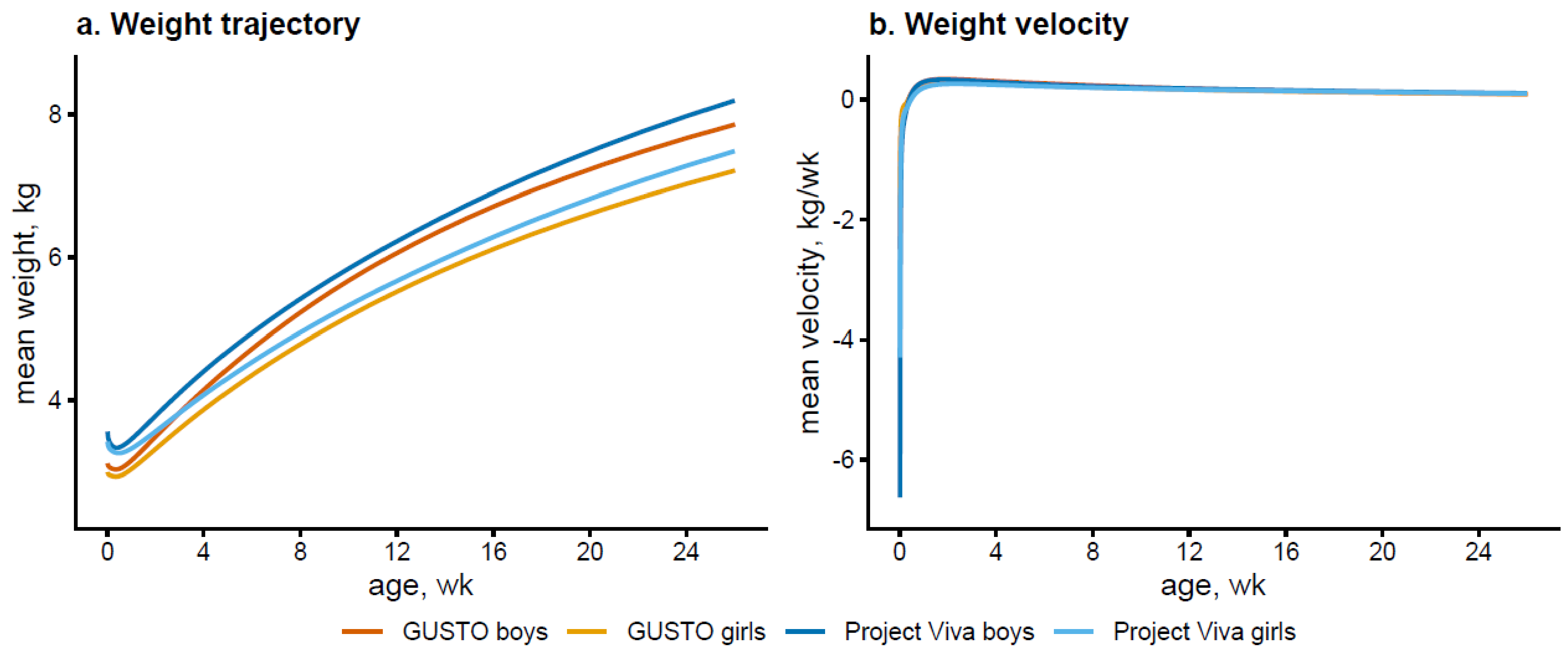
Study name	GUSTO	Project Viva
Region and country	Singapore	Massachusetts, USA
Particpant birth years	2009-2010	1999-2002
Particpant race/ethnicity	Chinese, Indian, Malay	White, Black, Hispanic, Asian, Mixed/Other

	GUSTO		Project Viva	
Participants, n	594	540	799	759
Age range, wk	0-26.0	0-26.0	0-26.0	0-26.0
Obs/particpant, median (range)	4 (2-6)	4 (2-6)	4 (2-9)	4 (2-8)

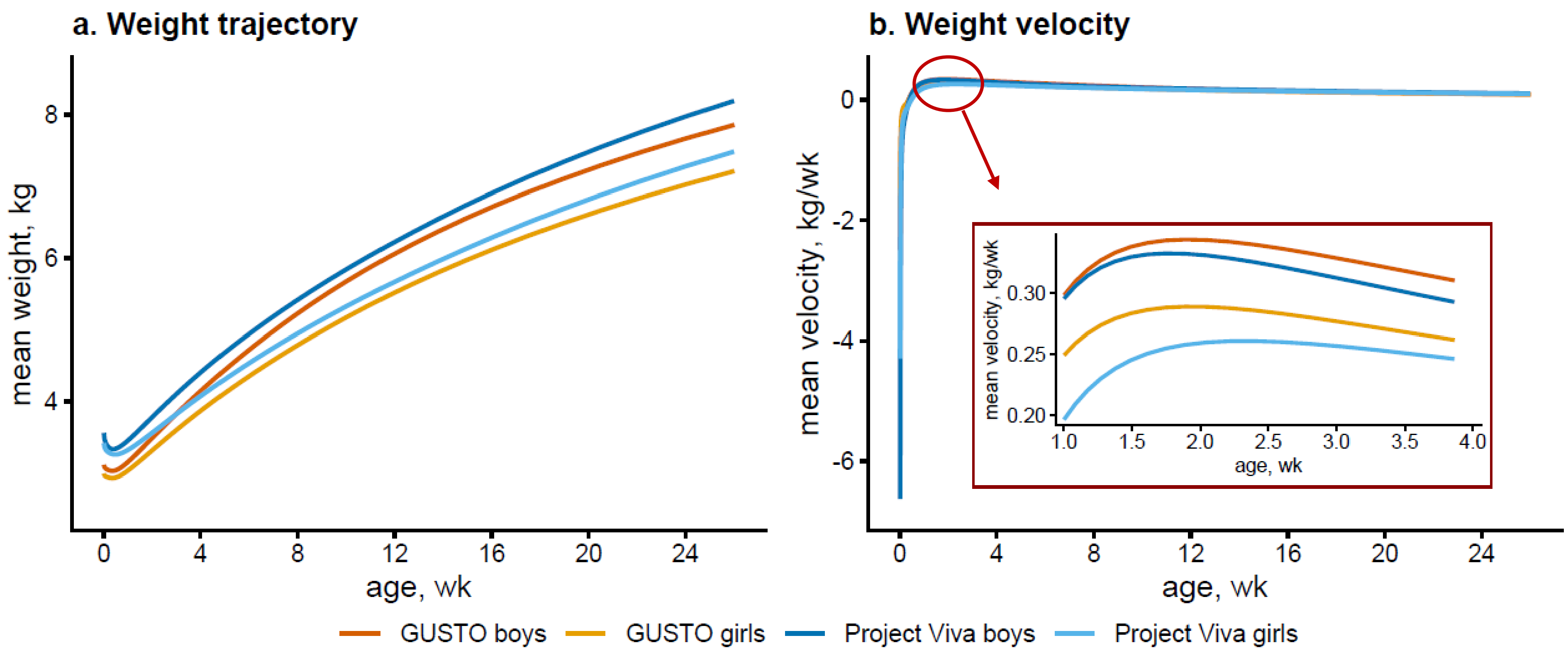
	df	RMSE (kg)	variance explained (%)
GUSTO boys	2	0.13	92.5
GUSTO girls	2	0.13	92.5
Project Viva boys	2	0.11	94.6
Project Viva girls	2	0.10	94.9

NCCOR

Mean distance and velocity curves: Birth to 6 months



Mean distance and velocity curves: Birth to 6 months



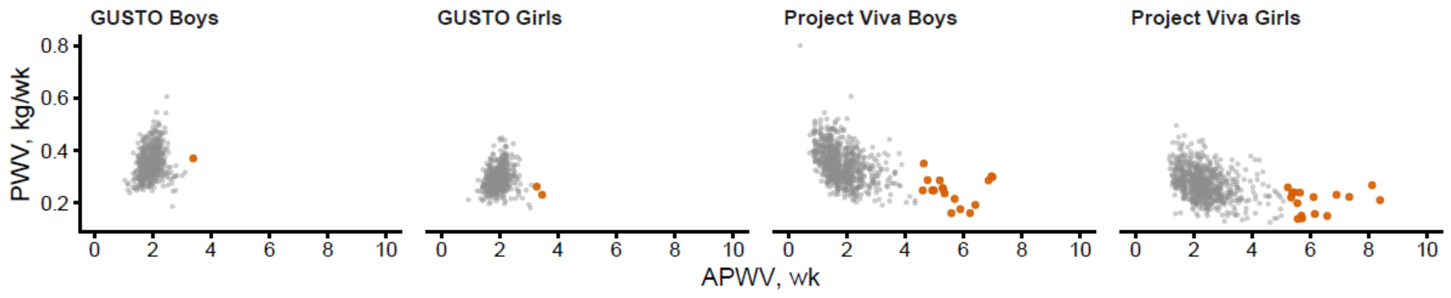
NCCOR

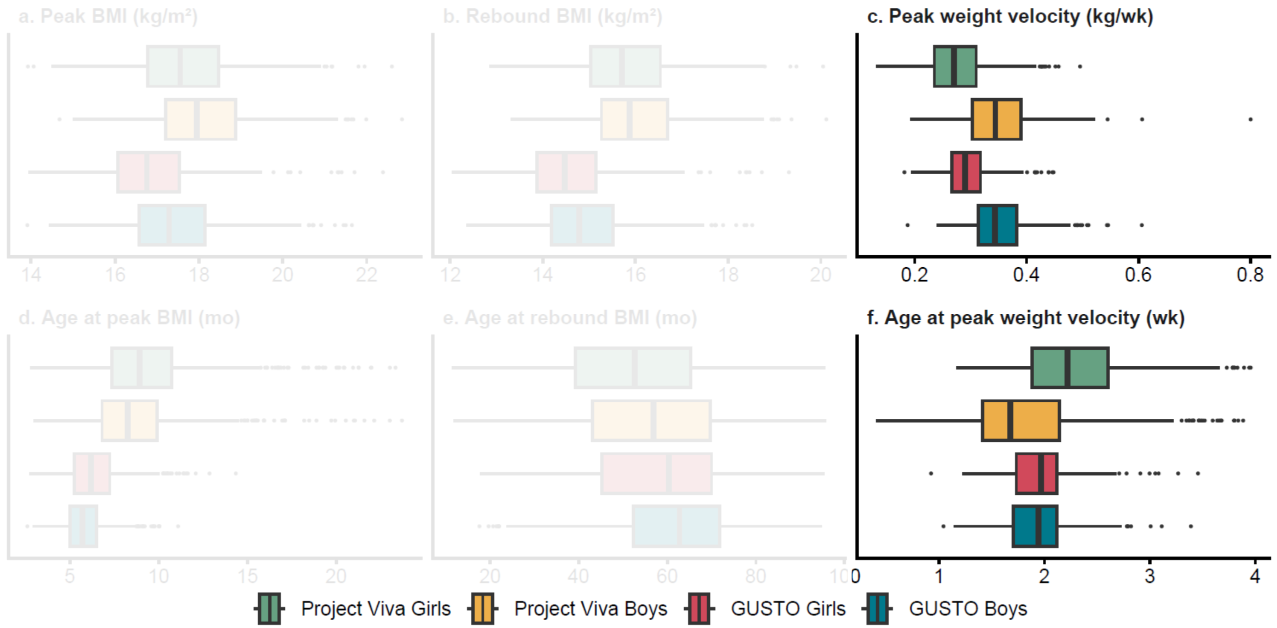
PWV and APWV

PWV estimated for all children (n=2692)

Boxplot shows data after removing 65 children with APWV > 4 weeks

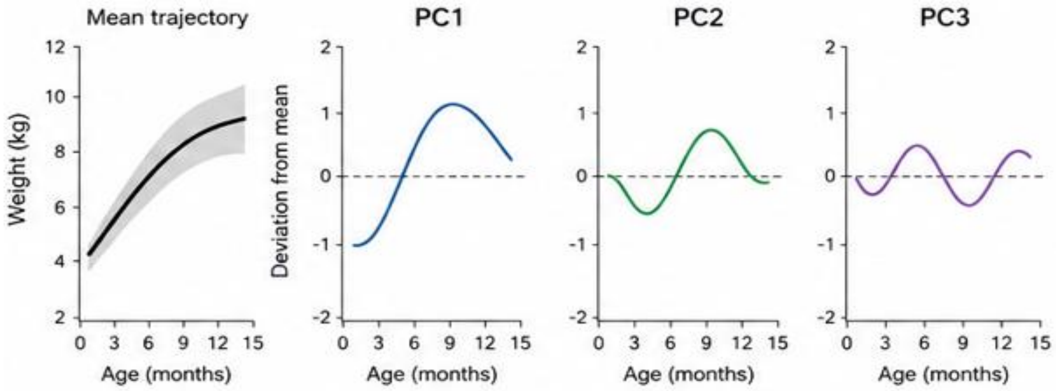
Fig: scatterplots of PWV agains APWV





Functional principal component analysis (FPCA)

- Models growth trajectory as a continuous function of age (function often smoothed via spline)
- Identifies main patterns of variation between individuals in their growth trajectories
- Decomposes individual trajectories into a mean trajectory + functional principal components (FPCs)
- Each FPC represents a distinct mode of variation in growth
- FPCs can be interpreted, and *can* represent differences in size, timing, velocity and shape of growth



Gertheiss, J. et al. (2024), Functional Data Analysis: An Introduction and Recent Developments. *Biometrical Journal*

FPCA example

626 Bangladesh children

Height measured every 3 months, up to 24 months

FPCA applied to height and HAZ

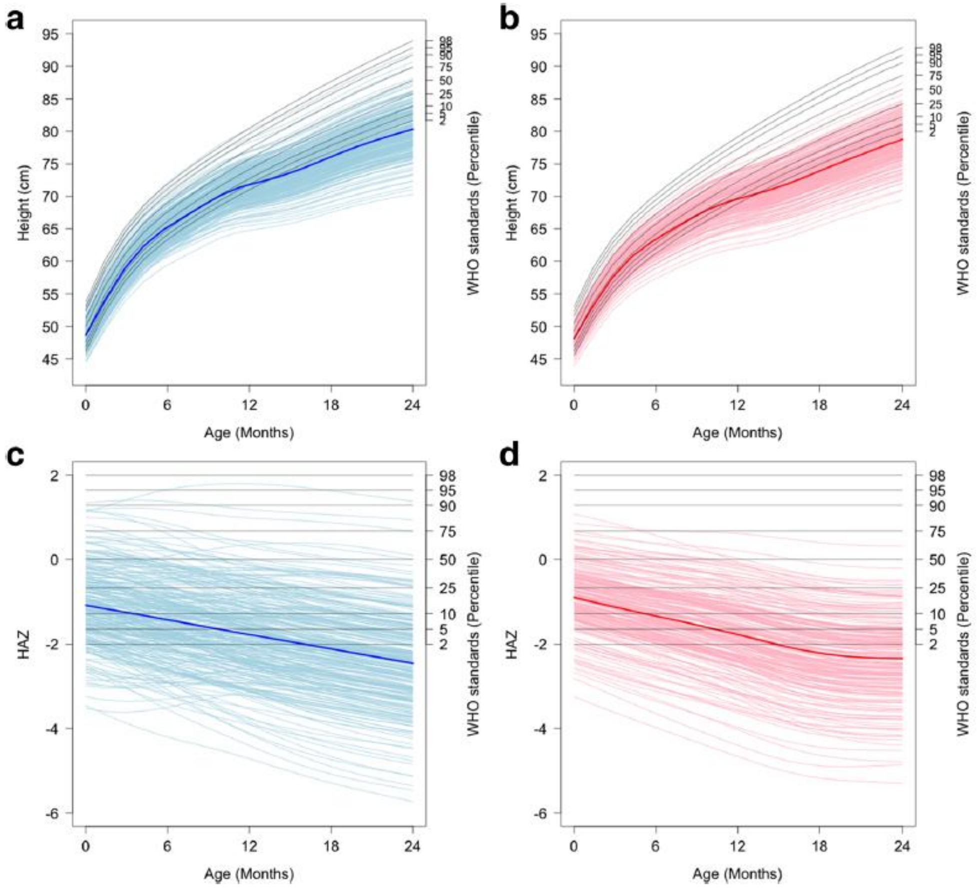


Fig. 1 Individually fitted curves of height (Panels **a** and **b**) and height-for-age z-score (HAZ) (Panels **c** and **d**) from birth to age 24 months for boys (*light blue*, n = 270) and girls (*pink*, n = 225). The mean *curves* are shown in *solid blue line* for boys and *red line* for girls. The World Health Organization (WHO) growth standards are shown as the *solid black lines* in Panels **a** and **b** and as *solid horizontal black lines* in Panels **c** and **d** for the WHO "pseudo children" who were growing along the WHO percentiles, separately for boys and girls

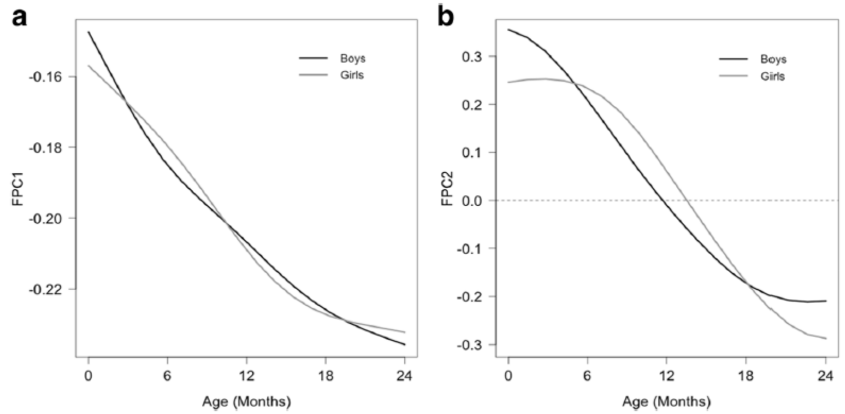


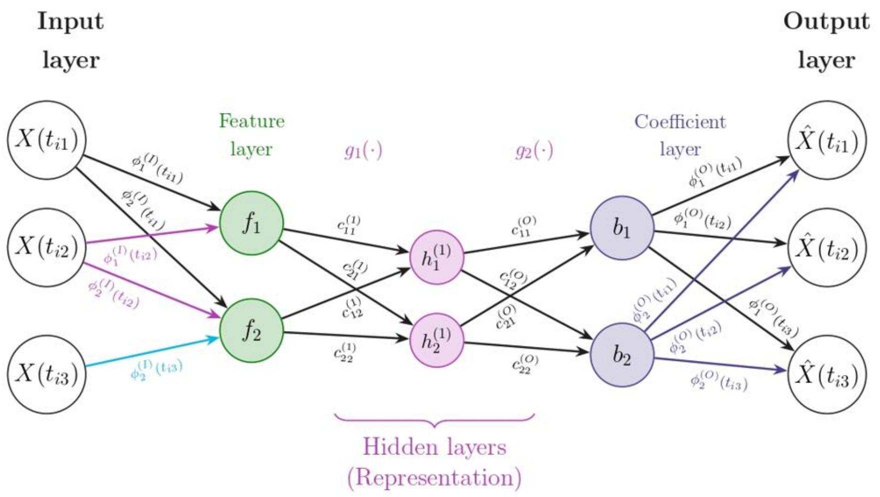
Fig. 2 The two leading functional principal components (FPCs) in boys (*black*) and girls (*grey*). The FPC1 in both boys and girls (Panel **a**) were *negative* and monotonically *decreased* over time, reflecting further deviation of individual growth patterns from the mean *curve*. The FPC2 (Panel **b**) changed signs approximately at 12 months for boys and 14 months for girls, indicating the changes in growth trajectory. For boys, the FPC1 and FPC2 accounted for 93 and 6% of the variation among fitted height-for-age z-score (HAZ) *curves*. For girls, the FPC1 and FPC2 accounted for 96 and 3% of the variation among fitted HAZ *curves*

Zhang Y et al. Characterizing early child growth patterns of height-for-age in an urban slum cohort of Bangladesh with functional principal component analysis. 2017 *BMC Pediatrics.*

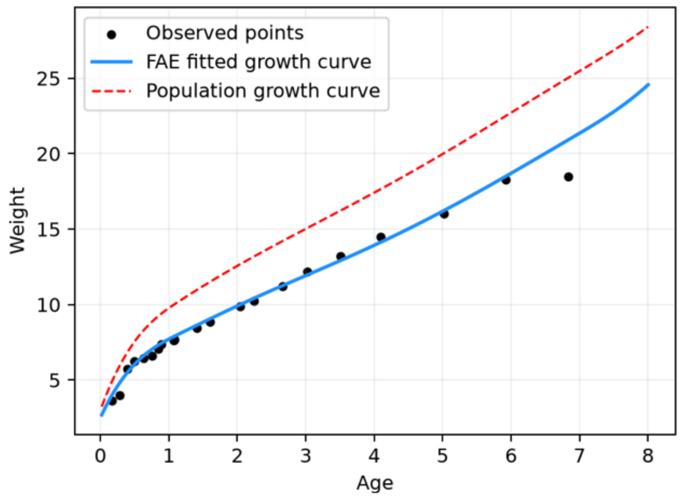
Ongoing work - functional autoencoder (FAE) for growth modelling

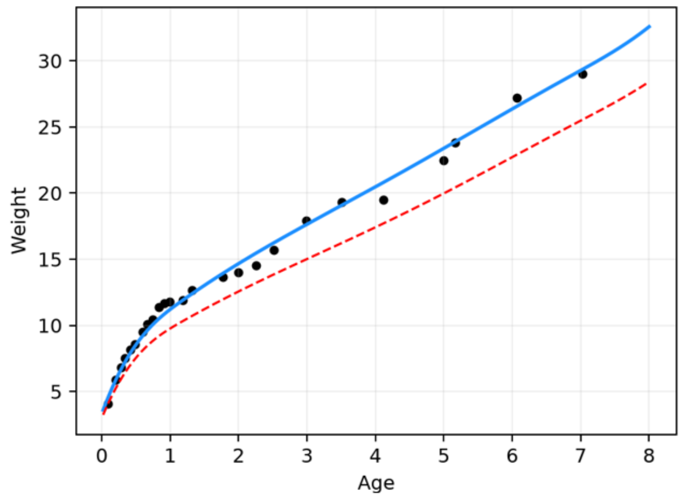
- FAE vs. FPCA
- Height, weight, and BMI up to 8 years

OSF PREREGISTRATION



A graphical representation of the FAE for discrete functional data. The model represented only have J = 3 time points, M^((I)) = 2 for the Feature layer, l=1 hidden layers with K^((1)) = K^((L)) = 1, and M^((O)) = 2 for the coefficient label





Hernandez MA et al. Modelling early life growth trajectories through representation learning with functional autoencoders. 2026. OSF Preregistration https://doi.org/10.17605/OSF.IO/TYPF4

Summary and research suggestions

- Lots of methods are available and useful for modelling growth curves, including others not covered here.
- Research needs include:
  - Methods comparison studies
  - Review of available methods
  - Tutorials/guides/recommendations on how to apply methods and analyst choices (e.g., age range, age/outcome transformation)
  - Better ways of studying growth velocity–better 2-stage approaches / directly modelling velocity / joint models