

RESEARCH ARTICLE

Functional principal component analysis for identifying multivariate patterns and archetypes of growth, and their association with long-term cognitive development

Kyunghee Han¹, Pantelis Z. Hadjipantelis¹, Jane-Ling Wang¹, Michael S. Kramer², Seungmi Yang², Richard M. Martin^{3,4}, Hans-Georg Müller^{1*}

1 Department of Statistics, University of California Davis, Davis, California, United States of America, **2** Departments of Pediatrics and of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada, **3** Bristol Medical School, Population Health Sciences, University of Bristol, Bristol, United Kingdom, **4** National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, United Kingdom

* hgmuller@ucdavis.edu



OPEN ACCESS

Citation: Han K, Hadjipantelis PZ, Wang J-L, Kramer MS, Yang S, Martin RM, et al. (2018) Functional principal component analysis for identifying multivariate patterns and archetypes of growth, and their association with long-term cognitive development. *PLoS ONE* 13(11): e0207073. <https://doi.org/10.1371/journal.pone.0207073>

Editor: Rhonda D. Szczesniak, Cincinnati Children's Hospital Medical Center, UNITED STATES

Received: June 13, 2018

Accepted: October 24, 2018

Published: November 12, 2018

Copyright: © 2018 Han et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The PROBIT data cannot be shared based on contractual terms in the participant consent agreement and inquiries can be made to Dr. Mourad Dahhou (Mourad.Dahhou@muhc.mcgill.ca), a representative of the data access committee for the PROBIT study in the McGill University Health Center.

Funding: This study was supported by the Bill & Melinda Gates Foundation (OPP1119700). The

Abstract

For longitudinal studies with multivariate observations, we propose statistical methods to identify clusters of archetypal subjects by using techniques from functional data analysis and to relate longitudinal patterns to outcomes. We demonstrate how this approach can be applied to examine associations between multiple time-varying exposures and subsequent health outcomes, where the former are recorded sparsely and irregularly in time, with emphasis on the utility of multiple longitudinal observations in the framework of dimension reduction techniques. In applications to children's growth data, we investigate archetypes of infant growth patterns and identify subgroups that are related to cognitive development in childhood. Specifically, "Stunting" and "Faltering" time-dynamic patterns of head circumference, body length and weight in the first 12 months are associated with lower levels of long-term cognitive development in comparison to "Generally Large" and "Catch-up" growth. Our findings provide evidence for the statistical association between multivariate growth patterns in infancy and long-term cognitive development.

Introduction

Objective of the study

The link between deficient growth in infancy and later life cognitive performance degradation has been widely accepted [1–3]. Stunting and faltering during infancy, or early childhood, are associated with reduced cognitive ability in later age performance [4, 5], and these growth patterns have been the subject of extensive investigation [6–9]. In most of the previous work, investigators have studied this association by examining single growth indicators, for example head circumference [10] or body weight [11]. In particular, [12] examined how cognitive

article contents are the sole responsibility of the authors and may not necessarily represent the official views of the Bill & Melinda Gates Foundation or other agencies that may have supported the primary data studies used in the present study.

Competing interests: The authors have declared that no competing interests exist.

development of children in Vietnam is associated with pre-defined growth features at age 1 year. While features such as stunting, underweight, wasting and small head circumference were examined, the previous analysis was based on one growth modality, for example body length or weight.

We propose a straightforward way to combine multiple growth indicators under a single framework. Our approach provides a comprehensive assessment of the potential risk in terms of cognitive development using longitudinal information from growth patterns of several growth modalities. Specifically, we demonstrate a data analysis procedure that combines measurements from three commonly recorded time-varying growth traits, head circumference, body length and body weight. We then identify growth patterns that can be associated with subsequently measured full-scale IQ (Wechsler abbreviated scale of intelligence, WASI). The simultaneous consideration of multiple trajectories is a main novel feature of our approach.

We also devise a simple method to learn archetypal growth patterns from data and examine their association with subsequent IQ outcomes. Our methods assess multiple growth indicators nonparametrically without the need of prior growth charts [13]. Using a functional data analysis (FDA) framework [14, 15], the proposed methodology combines multiple growth indicators and identifies data-driven clusters of infants according to their growth profiles. Recently, [16] and [17] considered quantile contour estimation of functional principal components (FPC) with emphasis on analysis for growth curves, but only with a single growth trajectory sample. In a related approach, [18] focused on finding subjective-specific warping functions to extract common features among multivariate growth traits. In contrast to existing approaches, we profile multiple growth patterns in terms of archetypal analysis [19–21], where we implicitly assume that extreme growth patterns can be used to represent individual growth curves in the sample through convex combination.

Our findings from applying the proposed methodology to the PROBIT growth study cohorts [22, 23] suggest that the proposed methodology is capable of identifying infant subgroups that differ in a statistically significant way in terms of the average level of associated IQs and thus can serve as a useful tool for identifying subgroups at risk of impaired cognitive development.

Data description

The data were collected as part of WHO's Promotion of Breastfeeding Intervention Trial (PROBIT) in the Republic of Belarus [22, 23]. They include growth measurements taken during the infancy of full term babies who weighed at least 2.5kg at birth among 17,045 total subjects. The physical traits recorded are head circumference, body length and body weight which were measured six times during the first year after birth. However, the sampling schedules were not strictly followed and some children did not have a full set of six measurements, the data are best characterized as irregularly sampled longitudinal observations and see Fig 1. For example, about 5.5% of the children have less than six measurements. For children with complete records, average measurement times were approximately 1.05 (0.12), 2.05 (0.13), 3.11 (0.24), 6.12 (0.31), 9.12 (0.34) and 12.10 (0.21) months after birth, where standard deviations at each visit are given in parentheses. We also refer to the design plot for measurement times during the first year in [13], which demonstrates the extent of sparsity and irregularity of the observation schedules over the first year after birth. The cognitive ability of children was assessed using the WASI score of the full-scale IQ measured at 6.5 years.

In addition to time-varying growth traits, a multitude of demographic covariates were also recorded. These covariates were used in later stages to control for potential confounding effects including socio-demographic factors. For example, children whose parents attended

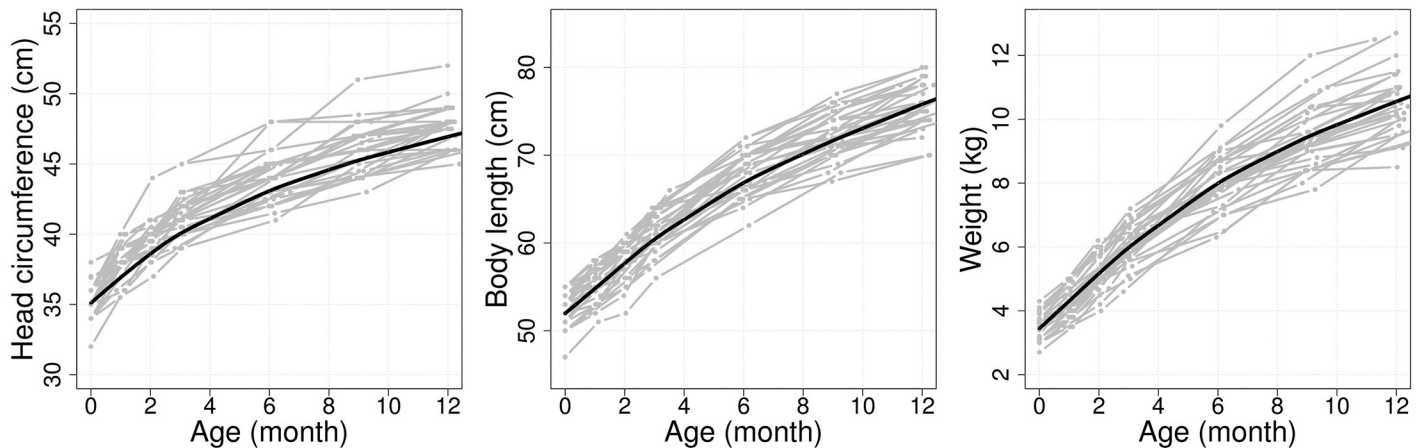


Fig 1. Irregular and sparse longitudinal observations from the PROBIT data. Head circumference (HC, left), body length (LN, middle) and weight (WT, right) are illustrated for a random selection of 30 subjects (gray) out of about 12,800 total children, along with estimated mean curves for each longitudinal trait (black solid lines).

<https://doi.org/10.1371/journal.pone.0207073.g001>

university have higher IQ measurements on average than children whose parents did not to complete high-school education. These covariates are known to affect later age IQ throughout a child’s infancy [24–27]. Other covariates included the sex of the child, maternal and paternal education levels, maternal and paternal age-at-birth, maternal smoking during pregnancy, duration of exclusive breast feeding and the hospital where the child was born. The sample size after preprocessing was $n = 12,809$ children, whose data were analyzed in our study. For details of data records and preprocessing, we refer to [13] and [22].

Methodology

Functional principal component analysis

Let $X_i(t)$ be a realization of a time-varying trait $X(t)$ for the i -th subject, $1 \leq i \leq n$, at each time point $t \in \mathcal{T}$. Assuming independent measurements between subjects and that the $X_i(t)$ have smooth trajectories over time t , we apply functional principal component analysis (FPCA) to decompose patterns of temporal variation. A basic feature of FPCA is that the time-varying trait of the i -th subject admits the Karhunen-Loève expansion [28–30]

$$X_i(t) = \mu(t) + \sum_{k \geq 1} \xi_{ik} \phi_k(t), \tag{1}$$

where $\mu(t) = EX(t)$, and the ξ_{ik} are uncorrelated random variables with mean zero and variance λ_k satisfying $\lambda_1 \geq \lambda_2 \geq \dots$. Here in Eq (1), the ξ_{ik} are the k -th FPC scores of the i -th subject, associated with the eigenfunction ϕ_k for all $k \geq 1$. For theoretical background on FPCA and related techniques, see [15, 31–36].

In longitudinal studies, however, measurements of time-varying traits are only available at N_i successive time points, say $t_{i1} < \dots < t_{iN_i}$, for the i -th subject. We note that the set of time points $\{t_{i1}, \dots, t_{iN_i}\}$ may differ among the n subjects and N_i may be small. FPCA for longitudinal data has been widely investigated [14, 37–40]. Specifically, [41, 42] proposed a technique to perform FPCA for sparse longitudinal data, based on principal components analysis through a conditional expectation (PACE) scheme. Specifically, we consider sparse and noisy

longitudinal observations $\tilde{X}_{ij} = X_i(t_{ij}) + \epsilon_{ij}$, instead of continuous and unperturbed observations of time-varying traits $X_i(t)$, where the ϵ_{ij} are independent mean zero measurement errors. By assuming that ξ_{ik} and ϵ_{ik} follow a joint normal distribution, the best linear predictors of the FPC scores ξ_{ik} are given by

$$\hat{\xi}_{ik} = \hat{\lambda}_k \hat{\phi}_{ik}^\top \hat{\Sigma}_{\tilde{X}_i}^{-1} (\tilde{\mathbf{X}}_i - \hat{\boldsymbol{\mu}}_i), \tag{2}$$

where $\tilde{\mathbf{X}}_i = (\tilde{X}_{i1}, \dots, \tilde{X}_{iN_i})^\top$ are longitudinal observations, $\hat{\boldsymbol{\mu}}_i = (\hat{\mu}(t_{i1}), \dots, \hat{\mu}(t_{iN_i}))^\top$ are the estimates of mean vectors of $E\tilde{\mathbf{X}}_i$, and $\hat{\Sigma}_{\tilde{X}_i}$ is the estimated $N_i \times N_i$ variance-covariance matrix of $\Sigma_{\tilde{X}_i}$ with (j, ℓ) -elements given by $\text{Cov}(\tilde{X}_{ij}, \tilde{X}_{i\ell})$. Also, $(\hat{\lambda}_k, \hat{\phi}_k(t))$, $k \geq 1$, are pairs of estimators for eigenvalues and eigenfunctions, which are the solutions of the following equations with respect to $(\lambda_k, \phi_k(t))$,

$$\int_{\mathcal{T}} G(s, t) \phi_k(s) ds = \lambda_k \phi_k(t) \quad (k = 1, 2, \dots),$$

$$\text{subject to } \lambda_k \geq \lambda_{k+1} \quad \text{and} \quad \int_{\mathcal{T}} \phi_k(t) \phi_\ell(t) dt = \begin{cases} 0 & (k \neq \ell) \\ 1 & (k = \ell) \end{cases}, \tag{3}$$

where $G(s, t) = \text{Cov}(X(s), X(t))$ is the auto-covariance function of X , so that we may write $\hat{\phi}_{ik} = (\hat{\phi}_k(t_{i1}), \dots, \hat{\phi}_k(t_{iN_i}))^\top$; see [43] and [15] for comprehensive overviews on FDA and recent developments in the interface between FPCA and longitudinal data.

Once we have estimated eigenfunctions $\hat{\phi}_k(t)$ through the PACE method in Eq (2), longitudinal patterns of $\tilde{\mathbf{X}}_i$ can be summarized by the corresponding FPC scores $\hat{\xi}_{ik}$. In fact, unobserved time-varying traits $X_i(t)$ can be reconstructed as $\hat{X}_i(t) = \hat{\mu}(t) + \sum_{k=1}^K \hat{\xi}_{ik} \hat{\phi}_k(t)$, followed by the representation in Eqs (1) and (3) with a cut-off value $K \geq 1$. The truncation point K can be chosen as the smallest value satisfying $\sum_{k=1}^K \hat{\lambda}_k / \sum_{\ell \geq 1} \hat{\lambda}_\ell \geq \kappa$ for a given $0 < \kappa < 1$, so that a fraction κ of variance is explained (FVE), see [15, 30]. The infinite dimensional functions $\tilde{\mathbf{X}}_i$ will then be represented by K -vectors $(\hat{\xi}_{i1}, \dots, \hat{\xi}_{iK})^\top$, which provides the required dimension reduction.

Identification of outlying subjects

To study archetypes in the multivariate data analysis framework, we cluster longitudinal observations $\tilde{\mathbf{X}}_i$ into subgroups based on trajectory patterns of reconstructed time-varying traits $X_i(t)$. Time-varying traits are recovered by the first few FPC scores with high fraction of variance explained (FVE). In practice, the first two FPC scores produce relatively clear discrimination of the data characteristics in many sparse and irregular longitudinal studies [41, 42, 44, 45]. As an exploratory illustration tool for outlier detection in multivariate data analysis, the bagplot [46] was introduced as a generalization of the univariate boxplot. In the bagplot, half-space location depth [47] is usually adopted so that the multivariate data points are ordered by an extended notion of univariate rank. The halfspace location depth $D(\mathbf{x}; \mathcal{X}_n)$ of a point $\mathbf{x} = (x_1, \dots, x_K)^\top \in \mathbb{R}^K$ over K -variate data $\mathcal{X}_n = \{\xi_i \in \mathbb{R}^K : 1 \leq i \leq n\}$ is defined by the smallest number of ξ_i contained in any halfspace with boundary line passing \mathbf{x} . Then, data points can be ordered by depth, that is $D(\xi_i; \mathcal{X}_n) \leq D(\xi_{i'}; \mathcal{X}_n)$, $1 \leq i \neq i' \leq n$. For modern concepts and related work on statistical depth, see also [48] and [49].

In this study, for the purpose of providing flexible inference based on sparse and irregularly observed functional and longitudinal data, we utilize the highest density region (HDR) as in [50] and [51]. We consider the $(1 - \alpha)$ -HDR for the K -variate data defined by

$$\mathcal{R}^\xi(1 - \alpha) = \{\mathbf{x} \in \mathbb{R}^K : f_\xi(\mathbf{x}) \geq f_\alpha\} \quad (0 < \alpha < 1), \tag{4}$$

where f_ξ is the joint density of a random vector $\xi = (\xi_1, \dots, \xi_K)^\top$ and

$$f_\alpha = \arg \max \left\{ y > 0 : \int_{\mathbb{R}^K} f_\xi(\mathbf{x}) \cdot \mathbb{I}(f_\xi(\mathbf{x}) \leq y) \, d\mathbf{x} \leq \alpha \right\} \tag{5}$$

in Eqs (4) and (5), respectively. Taking $\alpha = 0.05$ yields a support region where observations are expected to fall with at least 95% probability. We also note that the HDR captures the nature of the distribution of the data like location, scale, correlation and tail information in a flexible manner. [52] proposed a kernel-type estimator of $\mathcal{R}^\xi(1 - \alpha)$, where f_ξ and f_α are replaced by kernel density estimators, respectively. For example, one can use $\hat{f}_\xi(\mathbf{x}) = n^{-1} \sum_{i=1}^n \prod_{k=1}^K L_{h_k}(\xi_{ik} - x_k)$ with bandwidths $h_k > 0$, where $L_h(v) = L(v/h)/h$ is a scaled version of a baseline kernel L that is a probability density function with finite variance. The kernel estimator $\hat{\mathcal{R}}(1 - \alpha)$ of $\mathcal{R}(1 - \alpha)$ also enjoys level information of the joint density f_ξ . We identify $(100 \times \alpha)\%$ extreme subjects in a sample as those falling outside of $\hat{\mathcal{R}}(1 - \alpha)$. For densely observed functional data, recent studies have investigated several measures of functional outliers, such as band depth and extremal depth for functional data [53–55].

Joint feature extraction from multiple time-varying traits

In this subsection, we describe how we perform dimension reduction for multivariate longitudinal observations by employing the covariance structure between multiple traits. Using Eq (1) and the FVE method introduced in the previous subsection, let $X^{*,[j]}(t) = \sum_{k=1}^{K^{[j]}} \zeta_k^{[j]} \phi_k^{[j]}(t)$ be truncated versions of the original time-varying traits $X^{[j]}(t)$ using only the first $K^{[j]}$ eigenfunctions, where $\phi_k^{[j]}(t)$ is the k -th eigenfunction of the j -th trait, $1 \leq j \leq d$, $1 \leq k \leq K^{[j]}$, and let $\zeta_k^{[j]} = \xi_k^{[j]} / (\lambda_k^{[j]})^{1/2}$ denote the standardized k -th FPC score of the j -th longitudinal trait, respectively. Then the functional covariance structure among the truncated time-varying traits $(X^{*,[j]}(t); 1 \leq j \leq d)$ can be reduced to the variance-covariance matrix of $(\zeta_k^{[j]} : 1 \leq k \leq K^{[j]}, 1 \leq j \leq d)$. Indeed, $\text{Cov}(X^{*,[j]}(s), X^{*,[m]}(t)) = \sum_{k=1}^{K^{[j]}} \sum_{\ell=1}^{K^{[m]}} (\lambda_k^{[j]} \lambda_\ell^{[m]})^{1/2} \text{Cov}(\zeta_k^{[j]}, \zeta_\ell^{[m]}) \phi_k^{[j]}(s) \phi_\ell^{[m]}(t)$, $1 \leq j \neq m \leq d$. This suggests to apply conventional principal component analysis (PCA) on the vector of standardized marginal FPC scores $(\zeta_k^{[j]} : 1 \leq k \leq K^{[j]}, 1 \leq j \leq d)$. Then, time-varying associations among multiple time-varying traits can be reproduced by a few PC scores in this second analysis. This approach has strong connections with the joint functional analysis methods of multiple random processes [56–58], and we also refer to [59] for similar ideas in a recent study on relationships between univariate and multivariate functional principal component analyses.

Identifying subgroups for risk associated with outcomes

Our study aims to identify at-risk longitudinal growth patterns associated with undesirable outcomes. For this, we consider conditional density function of outcomes Y given a collection of multiple time-varying traits. Let $f_{Y|S}$ be the conditional density of Y given a collection S of principal components that are obtained from the principal component analysis of the marginal FPC scores $(\zeta_k^{[j]} : 1 \leq k \leq K^{[j]}, 1 \leq j \leq d)$. In this study, we suggest four clusters $(S_m; 1 \leq m \leq 4)$ of standardized principal components of the multiple traits based on the $(1 - \alpha)$ -HDR

method as follows:

$$\begin{aligned}
 S_1 &= \{\mathbf{Z} \notin \mathcal{R}(1 - \alpha) : |Z_1|/\rho_1^{1/2} > |Z_2|/\rho_2^{1/2}, Z_1 > 0\}, \\
 S_2 &= \{\mathbf{Z} \notin \mathcal{R}(1 - \alpha) : |Z_1|/\rho_1^{1/2} < |Z_2|/\rho_2^{1/2}, Z_2 > 0\}, \\
 S_3 &= \{\mathbf{Z} \notin \mathcal{R}(1 - \alpha) : |Z_1|/\rho_1^{1/2} > |Z_2|/\rho_2^{1/2}, Z_1 < 0\}, \\
 S_4 &= \{\mathbf{Z} \notin \mathcal{R}(1 - \alpha) : |Z_1|/\rho_1^{1/2} < |Z_2|/\rho_2^{1/2}, Z_2 < 0\},
 \end{aligned}
 \tag{6}$$

where $\mathbf{Z} = (Z_1, Z_2)^\top$ is a 2-vector consisting of the first two PC scores obtained from the PCA of $\zeta = (\zeta^{[1]}, \zeta^{[2]}, \zeta^{[3]})$ and $\mathcal{R}(1 - \alpha)$ is the $(1 - \alpha)$ -HDR of \mathbf{Z} as in Eq (4). Also, ρ_1 and ρ_2 are the eigenvalues associated with the first two PC scores Z_1 and Z_2 , respectively. We then examine distributional differences among $f_{Y|S_m}(\cdot|S_m)$, $1 \leq m \leq 4$, and quantify the distributional differences with analysis of variance (ANOVA).

For practical implementation, we use conditional kernel density estimators for $f_{Y|S}$, given by $\hat{f}_{Y|S}(y|\hat{S}_m) = |\hat{S}_m|^{-1} \sum_{i \in \hat{S}_m} K_h(Y_i - y)$ with a bandwidth $h > 0$. Here $|\hat{S}_m|$ equals the number of elements in \hat{S}_m , which are empirical clusters of Eq (6) defined by

$$\begin{aligned}
 \hat{S}_1 &= \{1 \leq i \leq n : \hat{\mathbf{Z}}_i \notin \hat{\mathcal{R}}(1 - \alpha), |\hat{Z}_{i1}|/\hat{\rho}_1^{1/2} > |\hat{Z}_{i2}|/\hat{\rho}_2^{1/2}, \hat{Z}_{i1} > 0\}, \\
 \hat{S}_2 &= \{1 \leq i \leq n : \hat{\mathbf{Z}}_i \notin \hat{\mathcal{R}}(1 - \alpha), |\hat{Z}_{i1}|/\hat{\rho}_1^{1/2} < |\hat{Z}_{i2}|/\hat{\rho}_2^{1/2}, \hat{Z}_{i2} > 0\}, \\
 \hat{S}_3 &= \{1 \leq i \leq n : \hat{\mathbf{Z}}_i \notin \hat{\mathcal{R}}(1 - \alpha), |\hat{Z}_{i1}|/\hat{\rho}_1^{1/2} > |\hat{Z}_{i2}|/\hat{\rho}_2^{1/2}, \hat{Z}_{i1} < 0\}, \\
 \hat{S}_4 &= \{1 \leq i \leq n : \hat{\mathbf{Z}}_i \notin \hat{\mathcal{R}}(1 - \alpha), |\hat{Z}_{i1}|/\hat{\rho}_1^{1/2} < |\hat{Z}_{i2}|/\hat{\rho}_2^{1/2}, \hat{Z}_{i2} < 0\},
 \end{aligned}
 \tag{7}$$

where the $\hat{\mathbf{Z}}_i$ are vectors of the first two PC scores from the PCA of $\hat{\zeta}_i = (\hat{\zeta}_i^{[1]}, \hat{\zeta}_i^{[2]}, \hat{\zeta}_i^{[3]})$ and $\hat{\mathcal{R}}(1 - \alpha)$ is the $(1 - \alpha)$ -HDR of $\hat{\mathbf{Z}}_i$ as defined in the previous subsection. Also, $\hat{\rho}_1$ and $\hat{\rho}_2$ are estimates of the eigenvalues associated with the first two PC scores, respectively.

Finally, we identify subgroups for risk associated with outcomes based on multiple comparison techniques. Once we find significant differences among different subgroups, post-hoc procedures can be applied to perform multiple comparisons and control for multiple testing, which then lends support to specify risk subgroups associated with outcomes. For example, Bonferroni or Benjamini-Hochberg [60] procedures can be applied for pairwise analysis and in the next section we adopt Tukey’s post-hoc analysis [61] as a multiple comparison procedure for testing mean differences between all pairs of groups. We also use the Kruskal-Wallis rank sum test [62] as a nonparametric procedure for one-way ANOVA, and the Tukey-Kramer test (or Nemenyi test) for pairwise comparisons.

Numerical illustrations

Simulation study

We demonstrate the finite sample performance of the proposed method to identify clusters of extreme subjects. For this purpose random trajectories $\mathbf{X} = (X^{[1]}, X^{[2]}, X^{[3]})$ were generated such that

$$X^{[j]}(t) = \mu_j(t) + \xi_1^{[j]} \phi_1^{[j]}(t) + \xi_2^{[j]} \phi_2^{[j]}(t), \quad t \in [0, 1],
 \tag{8}$$

for $1 \leq j \leq 3$, where the mean functions μ_j of $X^{[j]}$ were zero and we use the normalized Fourier basis $\phi_1^{[j]}(t) = \sqrt{2} \sin(2\pi t)$ and $\phi_2^{[j]}(t) = \sqrt{2} \cos(2\pi t)$ on the interval $[0, 1]$ for all $1 \leq j \leq 3$.

The FPC score vectors $\xi^{[j]} = (\xi_1^{[j]}, \xi_2^{[j]})^\top$ were generated with multivariate normal distributions

with mean zero, $sd(\xi^{[1]}) = \text{diag}(3.0, 2.5)$, $sd(\xi^{[2]}) = \text{diag}(3.0, 2.0)$ and $sd(\xi^{[3]}) = \text{diag}(3.0, 1.5)$. For simplicity, we considered a common cross-covariance matrix for $(\xi^{[j]}, \xi^{[k]})$, given by

$$\begin{pmatrix} \text{cov}(\xi_1^{[j]}, \xi_1^{[k]}) & \text{cov}(\xi_1^{[j]}, \xi_2^{[k]}) \\ \text{cov}(\xi_2^{[j]}, \xi_1^{[k]}) & \text{cov}(\xi_2^{[j]}, \xi_2^{[k]}) \end{pmatrix} = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix} \tag{9}$$

for $1 \leq j \neq k \leq 3$. Let (ρ_j, \mathbf{v}_j) be (eigenvalue/eigenvector) pairs of the variance-covariance matrix Σ_ξ of $\xi = (\xi^{[1]}, \xi^{[2]}, \xi^{[3]})$, satisfying $\rho_1 \geq \dots \geq \rho_6$, where $\det(\Sigma_\xi) \approx 154.8$. The first two eigenvectors are $\mathbf{v}_1 \approx (0.56, 0.14, 0.57, 0.11, 0.57, 0.09)^\top$ and $\mathbf{v}_2 \approx (-0.15, 0.75, -0.11, 0.51, -0.09, 0.38)^\top$, and the corresponding eigenvalues are $\rho_1 \approx 4.041$ and $\rho_2 \approx 3.068$ (FVEs are 26.94% and 20.46%, respectively). Then, a scalar response Y was generated by $Y = \beta_1 Z_1 + \beta_2 Z_2 + \epsilon$, where $Z_j = \mathbf{v}_j^\top \xi$, $\beta = (\beta_1, \beta_2)^\top = (0.4, 0.2)^\top$ and $\epsilon \sim N(0, 0.4^2)$.

From n random copies $(\mathbf{X}_i; 1 \leq i \leq n)$ of \mathbf{X} for $n = 1000$, we generated sparse and noisy observations $\tilde{X}_i^{[j]}(T_{ijk}) = X_i^{[j]}(T_{ijk}) + \epsilon_{ijk}$, $1 \leq k \leq N_{ij}$, where the N_{ij} are randomly chosen integers between 5 and 10, T_{ijk} are iid uniform random variables on $(0, 1)$ and ϵ_{ijk} are Gaussian measurement errors with mean zero and variance 0.1^2 , and N_{ij} , T_{ijk} and ϵ_{ijk} were generated independently. Let $\mathcal{Y}_n = \{(Y_i, \tilde{\mathbf{X}}_i^{[1]}, \tilde{\mathbf{X}}_i^{[2]}, \tilde{\mathbf{X}}_i^{[3]}) : 1 \leq i \leq n\}$ be the random sample generated as described above, where $\tilde{\mathbf{X}}_i^{[j]} = (\tilde{X}_i^{[j]}(T_{ijk}) : 1 \leq k \leq N_{ij})$ for $1 \leq j \leq 3$.

We also demonstrate the outcomes Y that are associated with extremes of the predictors \mathbf{Z} such that (high Z_1 , high Z_2), (high Z_1 , low Z_2), (low Z_1 , high Z_2) and (low Z_1 , low Z_2) entail different levels of response outcomes. For example, suppose that we have $\mathbf{Z}_1 = (1, 1)^\top$, $\mathbf{Z}_2 = (1, -1)^\top$, $\mathbf{Z}_3 = (-1, 1)^\top$ and $\mathbf{Z}_4 = (-1, -1)^\top$, then the corresponding conditional means of the response outcomes are 0.6, 0.2, -0.2 and -0.6 which may represent different risk levels of subgroups associated with outcomes. In Fig 2 we present an example of one i.i.d. sample from (Y, Z_1, Z_2) , where we demonstrate four archetypal clusters associated with different response outcomes in different colors. This simulation setting illustrates a simple case where archetypes of functional patterns are associated with response.

Since we only have sparse and noisy longitudinal observations $\tilde{\mathbf{X}}^{[j]}$ of $X^{[j]}$, we first estimate $E(\xi^{[j]} | \tilde{\mathbf{X}}^{[j]})$ for each j -th trajectory marginally by applying FPCA based on the conditional expectation technique (PACE) [41] to $(\tilde{\mathbf{X}}_i^{[j]} : 1 \leq i \leq n)$. For computation, we used the “fdapace” package in R [63], where $\{\hat{\xi}_i^{[j]} = (\hat{\xi}_{i1}^{[j]}, \hat{\xi}_{i2}^{[j]})^\top : 1 \leq i \leq n\}$ denotes estimates of $E(\xi_i^{[j]} | \tilde{\mathbf{X}}_i^{[j]})$, $1 \leq i \leq n$, obtained from the PACE algorithm. Implementing the proposed method described in the Methodology section, we obtain four clusters $\hat{S}_m^{[j]}$ based on the $(1 - \alpha)$ -HDR method and standardized PC scores of multiple traits as in Eq (7). We considered the performance of our proposed methodology for the identification of risk clusters in comparison to using the univariate traits separately. Similarly, we obtained the marginal four clusters $\hat{S}_m^{[j]}$ based on the 95%-HDR method analogously to the above and standardized the individual FPC scores of each j -th trait as follows:

$$\begin{aligned} \hat{S}_1^{[j]} &= \{1 \leq i \leq n : \hat{\xi}_i^{[j]} \notin \hat{\mathcal{R}}^{[j]}(1 - \alpha), |\hat{\xi}_{i1}^{[j]}|/(\hat{\lambda}_1^{[j]})^{1/2} > |\hat{\xi}_{i2}^{[j]}|/(\hat{\lambda}_2^{[j]})^{1/2}, \hat{\xi}_{i1}^{[j]} > 0\}, \\ \hat{S}_2^{[j]} &= \{1 \leq i \leq n : \hat{\xi}_i^{[j]} \notin \hat{\mathcal{R}}^{[j]}(1 - \alpha), |\hat{\xi}_{i1}^{[j]}|/(\hat{\lambda}_1^{[j]})^{1/2} < |\hat{\xi}_{i2}^{[j]}|/(\hat{\lambda}_2^{[j]})^{1/2}, \hat{\xi}_{i2}^{[j]} > 0\}, \\ \hat{S}_3^{[j]} &= \{1 \leq i \leq n : \hat{\xi}_i^{[j]} \notin \hat{\mathcal{R}}^{[j]}(1 - \alpha), |\hat{\xi}_{i1}^{[j]}|/(\hat{\lambda}_1^{[j]})^{1/2} > |\hat{\xi}_{i2}^{[j]}|/(\hat{\lambda}_2^{[j]})^{1/2}, \hat{\xi}_{i1}^{[j]} < 0\}, \\ \hat{S}_4^{[j]} &= \{1 \leq i \leq n : \hat{\xi}_i^{[j]} \notin \hat{\mathcal{R}}^{[j]}(1 - \alpha), |\hat{\xi}_{i1}^{[j]}|/(\hat{\lambda}_1^{[j]})^{1/2} < |\hat{\xi}_{i2}^{[j]}|/(\hat{\lambda}_2^{[j]})^{1/2}, \hat{\xi}_{i2}^{[j]} < 0\}, \end{aligned} \tag{10}$$

where $\hat{\mathcal{R}}^{[j]}(1 - \alpha)$ is the $(1 - \alpha)$ -HDR of $\hat{\xi}_i^{[j]}$.

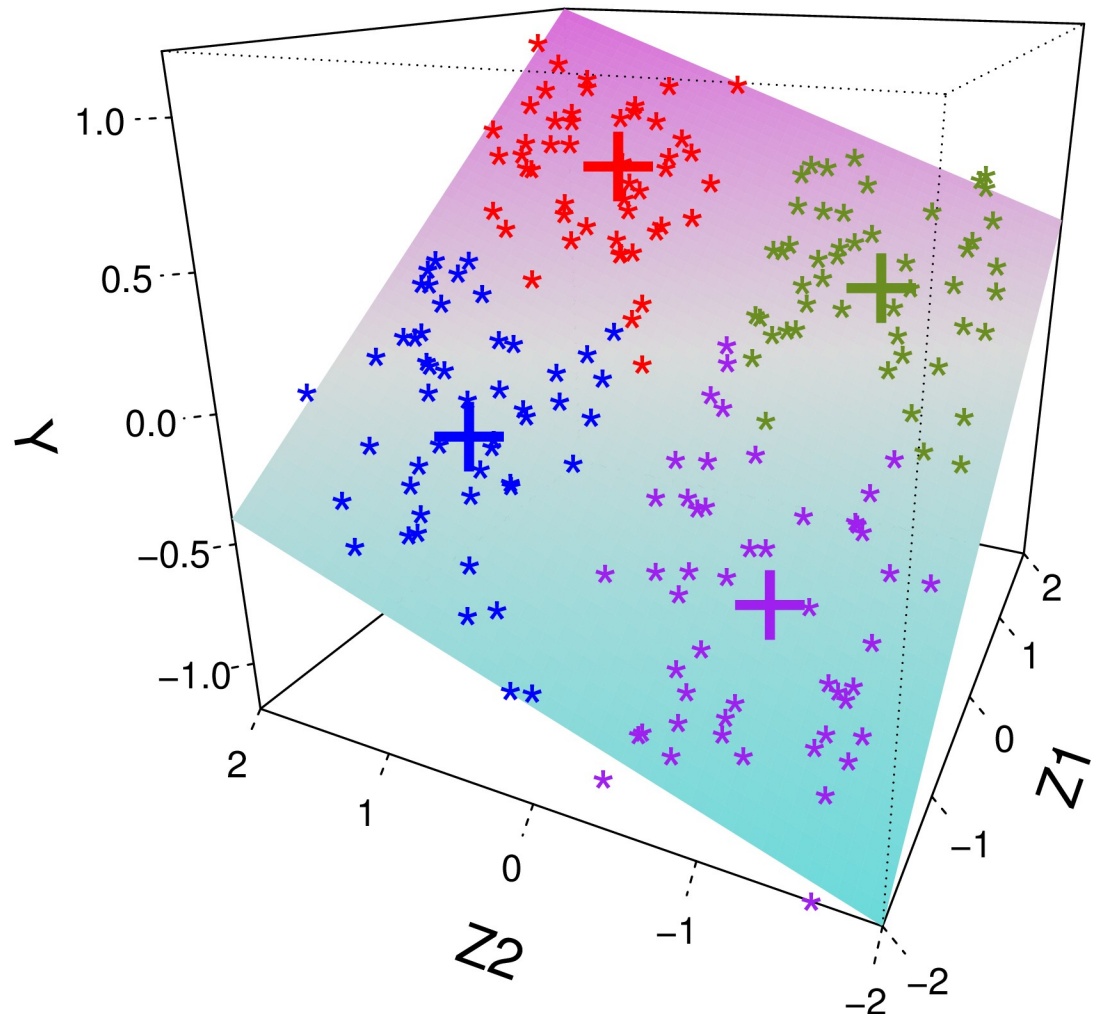


Fig 2. Example for visualization of observations from simulation. The proposed method identifies clusters associated with response outcomes Y characterized by archetypal covariate levels $Z = (Z_1, Z_2)$, for example (high Z_1 , high Z_2), (high Z_1 , low Z_2), (low Z_1 , high Z_2) and (low Z_1 , low Z_2), which are symbolized by red, purple, green and blue points, respectively, where the crosses denote the cluster centers. The surface demonstrates the conditional mean response when regressing Y on $Z = (Z_1, Z_2)$ for $n = 200$ data points.

<https://doi.org/10.1371/journal.pone.0207073.g002>

We report the simulation results in Table 1, where the numbers of joint extreme trajectory clusters associated with outcomes obtained from 1000 Monte Carlo repetitions with sample size $n = 1000$ are shown for $\alpha = 0.05$. Tukey’s post-hoc multiple comparison was employed to determine how many associated clusters exist at each Monte Carlo run. That is, at each repetition, we counted the subgroups which are completely separated by Tukey’s post-hoc analysis. By comparing conditional mean differences of outcomes between the four extreme clusters, we found that the proposed method identified more risk clusters than the marginal methods which detected two clusters on average for all cases. The joint method identified the three or four of the archetype clusters which depict (high Z_1 , high Z_2), (high Z_1 , low Z_2), (low Z_1 , high Z_2) and (low Z_1 , low Z_2) up to 90.4% (= 59.8% + 30.6%). This result supports the use of multiple trajectories instead of a single trajectory when identifying archetypes of risk sets. This applies even as the first two PC scores have less than 50% FVE, as in this simulation example.

Table 1. Simulation results for identification of at-risk multiple trajectories clusters associated with response outcomes.

Number of risk clusters associated with outcomes	Marginal method			Joint method PC-FPC
	$j = 1$	$j = 2$	$j = 3$	
< 2	3.6%	5.0%	4.7%	0.0%
2	62.3%	83.0%	84.0%	9.6%
3	31.7%	11.5%	10.8%	59.8%
4	2.4%	0.5%	0.5%	30.6%

For 1000 Monte Carlo (MC) repetitions with sample size $n = 1000$, the numbers of risk clusters were identified by analysis of variance (ANOVA) and Tukey’s multiple comparisons with a family-wise significance level 0.05. At each repetition, we counted subgroups completely separated by Tukey’s post-hoc analysis. For example, we identify two clusters if all subgroups included in a cluster show significant differences in pairwise comparison (family-wise significance level 0.05) against the other cluster members. Percentages in each column of the table demonstrate how many clusters are detected through 1000 MC repetitions. For the comparison with the marginal method, we applied the same procedure, using the marginal trajectory information only for $j = 1, 2, 3$, respectively.

<https://doi.org/10.1371/journal.pone.0207073.t001>

Analysis of PROBIT data

Marginal analysis for longitudinal measurements of growth traits. PROBIT contains three main time-varying traits; head circumference (HC), body length (LN) and weight (WT). For the marginal FPCA of these three variables, we applied the PACE technique introduced in the Methodology section, since we only have sparse and irregular observations available. As in the simulation study, we also used the “*fdapace*” package in R [63]. Auto-covariance functions of each time-varying trait were reconstructed by the first two eigenfunctions. The fractions of variance explained (FVEs) were 97.70%, 96.92% and 98.14% for HC, LN and WT, respectively. See Fig 3 for illustrations of estimated auto-covariance functions and eigenfunctions. The first and second eigenfunctions can be regarded as “General growth” and “Growth acceleration”, respectively. Based on the observed high FVE coverages, we assume in the rest of the paper that these two qualitative features carry information about the longitudinal patterns of time-varying growth traits in the PROBIT data.

The marginal analysis of outlying subgroups was performed with a $(1 - \alpha)$ -high density region (HDR) with $\alpha = 0.05$. Subjects were classified as “normal” if they belonged to the 95% support region in the HDR criterion as in Eq (4), while outlying subjects were classified as 5% extreme cases falling outside the HDR criterion. In this study, we considered four exclusive subgroups as in Eq (10), where the trait index $[j]$, $1 \leq j \leq 3$, stands for HC, LN and WT, respectively. The four outlying subgroups correspond to distinctive growth patterns, which can be labeled as “Generally Large”, “Catch-up”, “Stunting” and “Faltering” (Fig 4). We found that the outlying patterns were discordant across traits. For example, subjects who were classified into the generally large head circumference subgroup could be normal for body length or weight, and vice versa. Moreover, as there are 4^3 -combinations of subgroups entailed by the marginal analysis, it is difficult to associate all these multiple trajectory patterns with the response of interest, which is IQ at 6.5 years.

Potential risk subgroups for cognitive development. We constructed joint outlying subgroups of multiple time-varying traits based on the HDR method and standardized PC scores of multiple traits ($\hat{\zeta}_i^{[j]} = \hat{\zeta}_{ik}^{[j]} / (\hat{\lambda}_k^{[j]})^{1/2} : 1 \leq k \leq 2, 1 \leq j \leq 3$) as described in Eq (7). Principal component analysis results for the six FPC scores are presented in Table 2, where the first and second FPC stand for scores of general growth and growth acceleration, respectively. We found that these two features were captured in the first two PC loadings. In this study, we

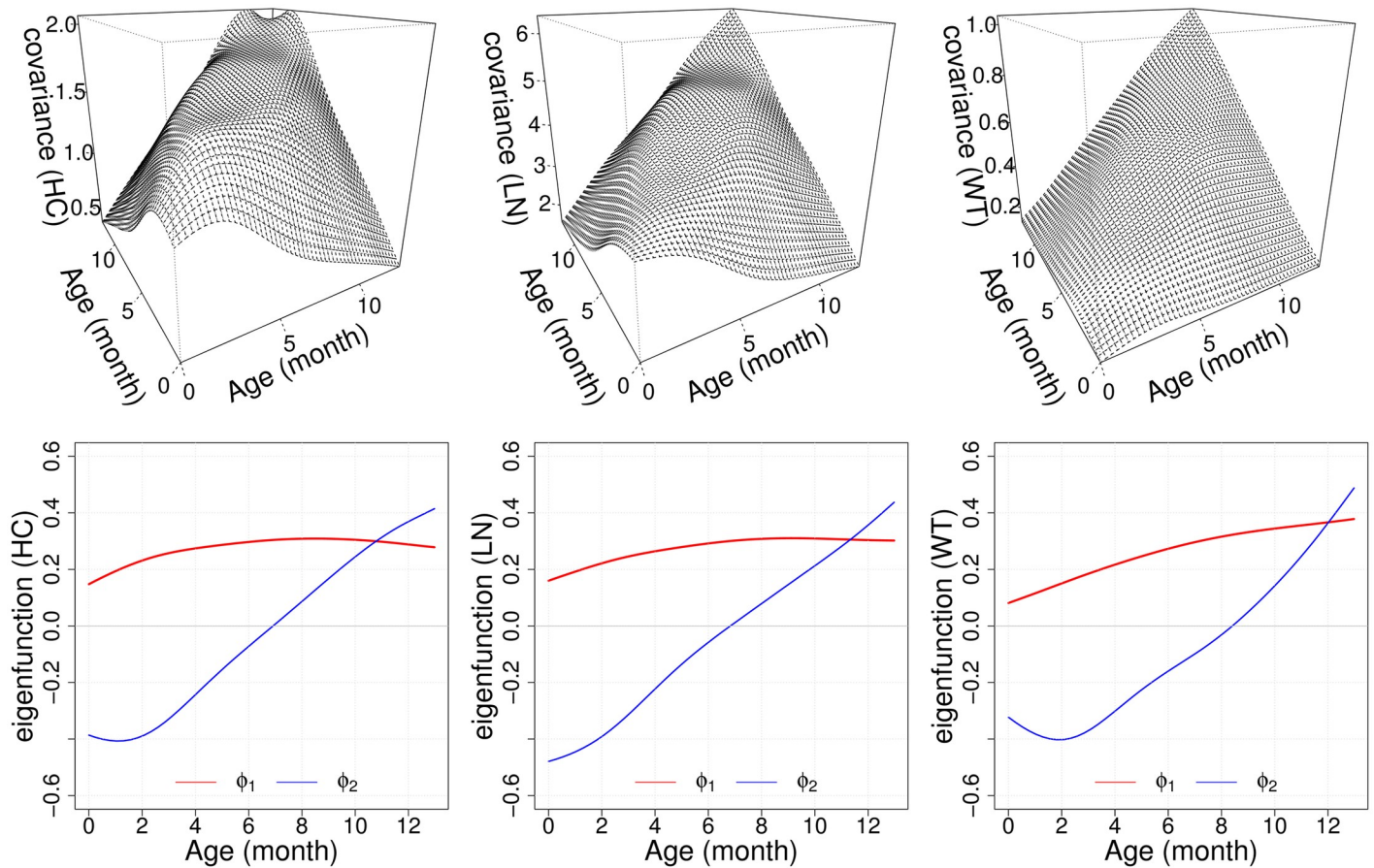


Fig 3. Estimated auto-covariance functions and eigenfunctions. Estimated auto-covariance functions $G(s, t)$ (top) and the corresponding first two eigenfunctions $\phi_1(t)$ and $\phi_2(t)$ (bottom) as in Eq (3) for head circumference (HC, left), body length (LN, middle) and weight (WT, right), respectively. Eigenfunctions represent the qualitative factors “General growth” (red) and “Growth acceleration” (blue). The cumulative fractions of variation explained (FVE) of the first two components are 97.70%, 96.92% and 98.14% for HC, LN and WT, respectively.

<https://doi.org/10.1371/journal.pone.0207073.g003>

focus on the first two PC scores as they explain more than 95% of the variation for each of the three modalities [41, 42, 45].

On the other hand, socio-economic factors affect childhood intelligence in ways that are not reflected in the FPCA of time-varying traits [26, 27]. To avoid confounding effects by socio-economic variables, we used a linear mixed effects model to reduce the influence of the socio-economic indicators. Hospital information was treated as a random effect as it is related to the random clustered design of the PROBIT study. We used the residuals of the linear mixed effects model and marginalized the effect of the potentially confounding variables considered above. For details of the data preprocessing, we refer to [22] and [13].

In Fig 5, we find that conditional densities of IQ measured at 6.5 years, given the joint outlying subgroups, exhibited different distributional behaviors. The four subgroups were constructed by a principal component analysis of the standardized six FPC scores for HC, LN and WT. The significance of the group mean differences was examined by one-way ANOVA (p-value = 0.002), and also by the Kruskal-Wallis rank sum test [62], a nonparametric procedure for one-way ANOVA (p-value < 0.001), so that the results were qualitatively the same. For a more detailed comparison, we performed post-hoc analysis with Tukey’s multiple comparison procedure. As shown in Table 3 and Fig 5, we found significant mean differences among

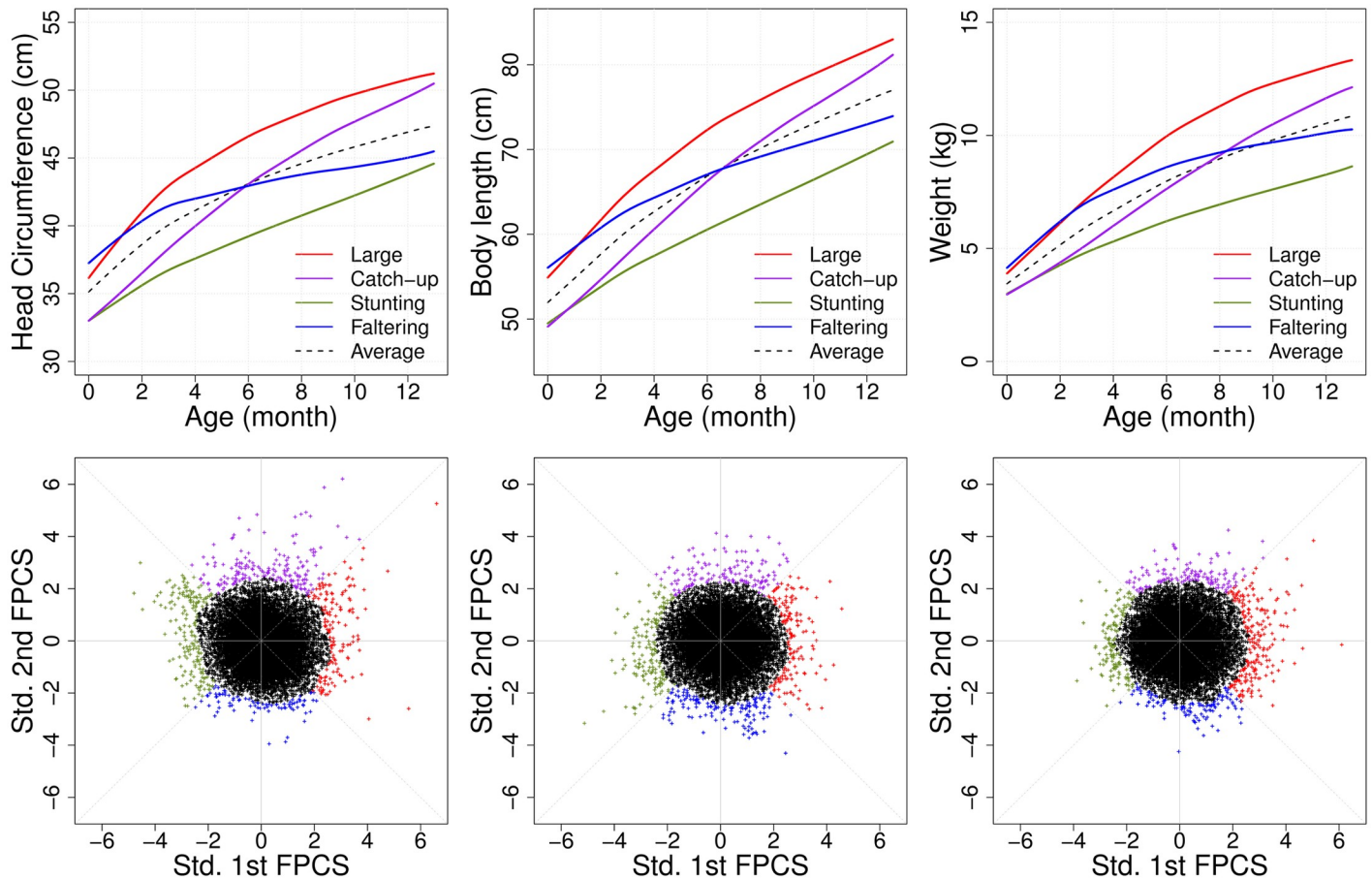


Fig 4. Extreme functional patterns from marginal analysis. Head circumference (HC, left), body length (LN, middle) and weight (WT, right) traits, respectively. Four outlying clusters (falling into the smallest 5% of the bivariate density) are demonstrated with respective different colors, with estimated mean curves corresponding to the four outlying subgroups, respectively, which represent the four qualitative longitudinal growth patterns: “Generally Large” (red), “Catch-up” (purple), “Stunting” (green) and “Faltering” (blue) (top panels). The corresponding scatterplots for the first two functional principal component scores (bottom).

<https://doi.org/10.1371/journal.pone.0207073.g004>

Table 2. Principal component analysis of functional principal components.

qualitative feature of joint FPCs	marginal FPC factor	PC loadings					
		PC1	PC2	PC3	PC4	PC5	PC6
General growth	HC-FPC1	0.502	0.155	0.457	0.690	0.044	0.191
	LN-FPC1	0.552	0.198	-0.289	-0.430	0.243	0.574
	WT-FPC1	0.527	0.331	-0.109	-0.182	-0.383	-0.649
Growth acceleration	HC-FPC2	-0.216	0.517	0.688	-0.422	0.188	-0.020
	LN-FPC2	-0.185	0.546	-0.432	0.324	0.570	-0.227
	WT-FPC2	-0.291	0.512	-0.190	0.154	-0.657	0.402
partial FVE		0.382	0.261	0.122	0.094	0.088	0.053
cumulative FVE		0.382	0.643	0.765	0.859	0.947	1.000

Principal component analysis for the variance-covariance matrix of the first two marginal functional principal component scores (FPCs) for head circumference (HC), body length (LN) and weight (WT).

<https://doi.org/10.1371/journal.pone.0207073.t002>

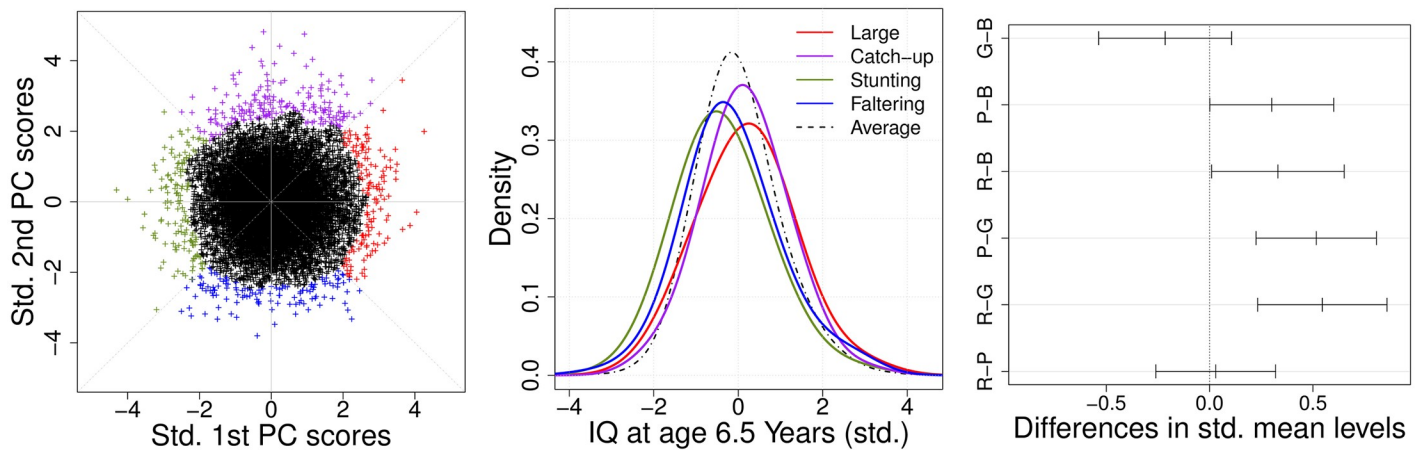


Fig 5. Extreme subgroup identification from the proposed method. (Left panel) Scatterplot of 95% high density region clustering from principal component analysis for head circumference, body length and weight and (Middle panel) the corresponding conditional kernel density estimates for long-term intelligence for each subgroup. Here, qualitative growth patterns include “Generally Large” (red), “Catch-up” (purple), “Stunting” (green) and “Faltering” (blue), whereas the black dashed line represents the “normal” subgroup that consists of subjects who do not belong to the four outlying subgroups. (Right panel) Tukey’s multiple comparisons of mean differences for standardized IQ along outlying subgroups are demonstrated by family-wise 95% confidence intervals, where we label the four subgroups as R (red, “Generally Large”), P (purple, “Catch-up”), G (green, “Stunting”) and B (blue, “Faltering”).

<https://doi.org/10.1371/journal.pone.0207073.g005>

outlying subgroups in a family-wise 5%-level test. “Stunting” and “Faltering” were associated with higher risk in comparison with the “Generally Large” and “Catch-up” subgroups. Similar results were obtained by using the nonparametric procedure for pairwise comparisons of the Tukey-Kramer test. We also applied several multiple comparison techniques such as Bonferroni and Benjamini-Hochberg methods for post-hoc analysis and similar results were obtained by controlling false discovery rate (FDR) at 5%. For example, we found that “Stunting” and “Faltering” were associated with higher risk in comparison with the “Generally Large” and “Catch-up” subgroups after application of both procedures. The results were suggestive of higher risk for “Faltering” versus “Catch-up” (p-value = 0.060 after Bonferroni correction).

We close this section with a short remark on Figs 6 and 7 which presents the result of the marginal analysis as described in the Simulation study section. In contrast to the proposed joint method, we found that the marginal procedures may not effectively detect risk growth patterns associated with long-term IQ outcomes. From the post-hoc analysis, “Generally Large” and “Stunting” subgroups had significantly different IQ performance for head circumference (p-value = 0.002), body length (p-value = 0.003) and weight (p-value < 0.001), but no significant difference was found between other subgroups such as “Catch-up” or “Faltering” for head circumference and body length. We note that “Stunting” for one of the marginal components may not be an at-risk growth pattern associated with IQ development compared to

Table 3. One-way ANOVA for subgroup detection.

variation	df	sum of sq.	mean sq.
subgroup	3	27.002	9.001
residuals	598	631.902	1.057
total	601	658.904	
F-value = 8.518		p-value < 0.001	

One-way ANOVA result provides evidence for differences in group means of long-term IQ for the four outlying subgroups “Generally Large”, “Catch-up”, “Stunting” and “Faltering”.

<https://doi.org/10.1371/journal.pone.0207073.t003>

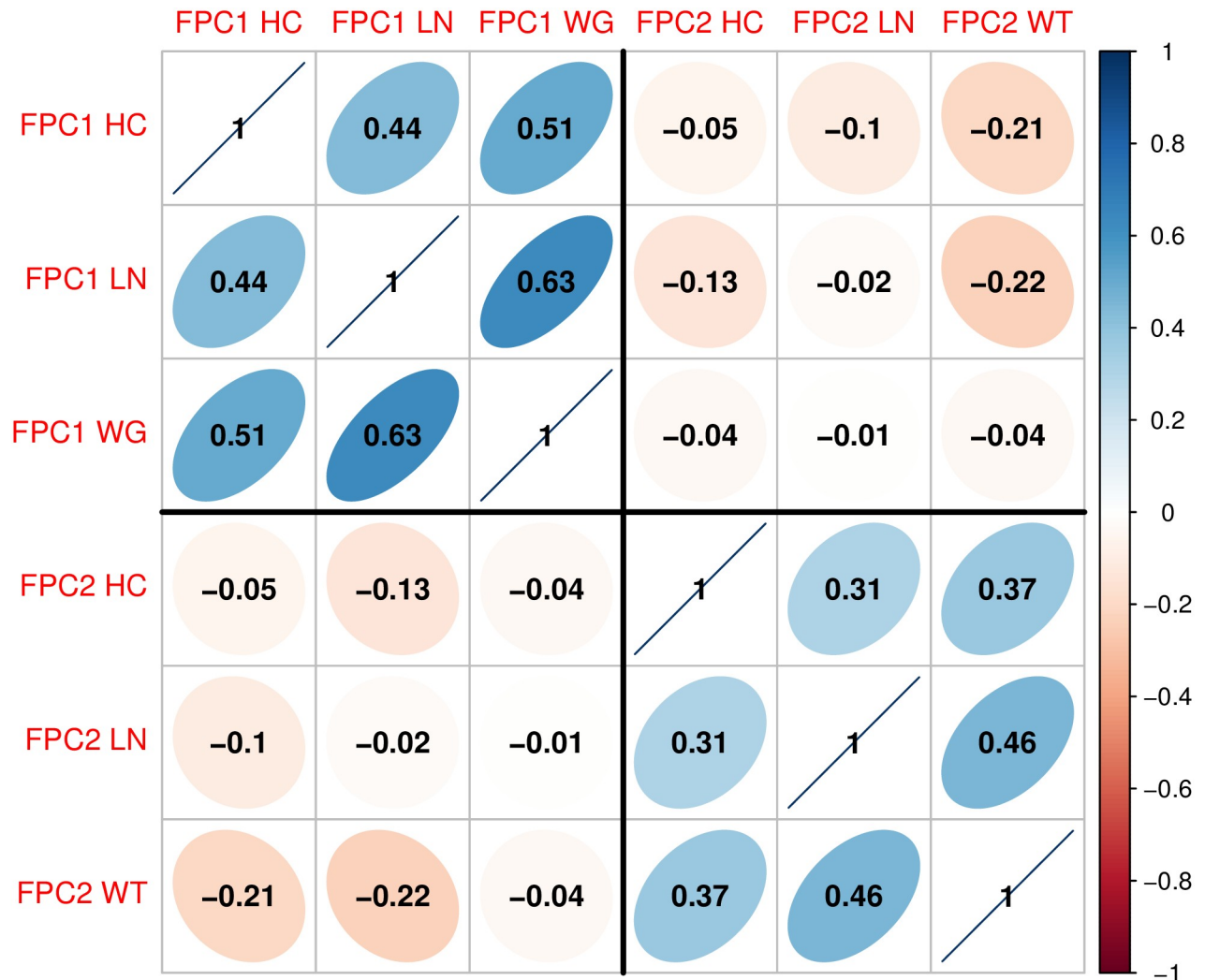


Fig 6. Correlation plot. Correlations between the first two functional principal component (FPC) scores of head circumference (HC), body length (LN) and weight (WT). FPC scores among the first and second components have positive correlations, respectively, which suggests to combine the three growth features linearly with PC loadings as in Table 2.

<https://doi.org/10.1371/journal.pone.0207073.g006>

the other subgroups. Moreover, “Faltering” for weight showed higher IQ performances than the “Stunting” subgroup (p-values < 0.001), while failure to thrive in infancy, defined as weight faltering in the first 9 months of life, was previously found to be associated with persistent deficits in intellectual development when measured at 8 years [64]. These results suggest that the combination of multiple growth patterns can indeed be beneficial for identifying risk subgroups associated with IQ outcomes.

Discussion

This paper outlines a statistical framework for exploring multivariate functional patterns deduced from sparsely and irregularly sampled longitudinal data and their association with long-term outcomes. For the joint analysis of children’s growth and IQ in the PROBIT data, we propose a straightforward way to combine multiple growth indicators under a single framework. We extract multiple growth features jointly, by using standard multivariate analysis of

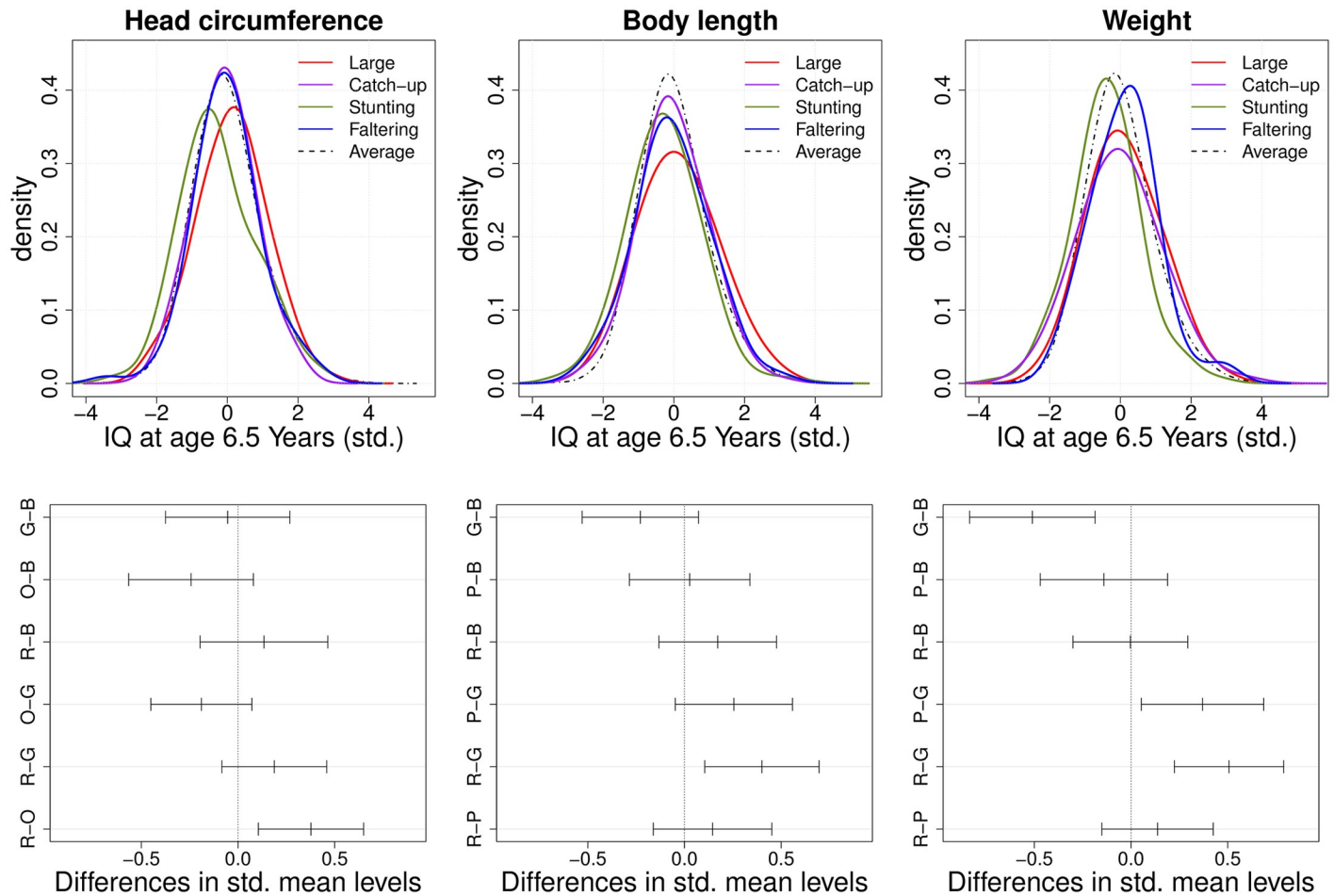


Fig 7. Marginal subgroup identification. (Top panels) Conditional kernel density estimators and (bottom panels) illustrations of Tukey’s multiple comparisons for standardized IQ along outlying subgroups. As in Fig 5, the first two functional principal component (FPC) scores are used to construct subgroups for head circumference (left), body length (middle) and weight (right), labeled R (red, “Generally Large”), P (purple, “Catch-up”), G (green, “Stunting”) and B (blue, “Faltering”).

<https://doi.org/10.1371/journal.pone.0207073.g007>

the functional principal components. The major modes of growth variation are then represented at the subject level and we can thus profile outlying multiple growth patterns, which can be considered as archetypes of growth.

The focus of this paper is how to combine multivariate functional data to identify extremal curve patterns and associate these features with responses. One may consider an alternative application of multivariate functional principal component analysis as in [58, 59, 65, 66] or functional ANOVA [67–69] as alternatives. However, for all approaches it is critical how to determine outlying and extremal patterns jointly from multiple functional data, and the high-density region (HDR) method to detect outlying functional principal components that we adopt here is a natural extension of similar nonparametric approaches in multivariate data analysis. Recently several related studies have been introduced for functional outlier detection [54, 70, 71] but it still remains an open problem how to combine these with other methodologies such as clustering and archetypal analysis.

In the PROBIT growth data analysis, we identified four archetypal subgroups of infant growth patterns, namely “Stunting”, “Faltering”, “Generally Large” and “Catch-up”. In addition we also found that covariance structures have marginally similar patterns across the

functional traits considered; head-circumference, body length and weight. According to our analysis, subgroups corresponding to “Stunting” and “Faltering” in the infant period had lower downstream IQ compared to “Generally Large” and “Catch-up” subgroups. This finding is supported by previous studies that link deficient infant growth and later life cognitive performance degradation.

It is worth mentioning that single growth indicators were not found to be associated with risk of lowered IQ, and the marginal analysis of single growth traits did not produce informative results in the PROBIT analysis (See Fig 7). Also, there is a possibility that the absence of any measure of cognitive ability during infancy in the data could be explained by reverse causality, namely, poor cognitive function in infancy may have led to worse dietary intake. Both may have been consequences of poor parenting or unmeasured insults during pregnancy or infancy. The proposed methods are not limited to specific data structures, such as growth data, but can be applied to many other kinds of longitudinal data as well, whenever a downstream outcome is of interest.

Acknowledgments

This study was supported by the Bill & Melinda Gates Foundation (OPP1119700). The article contents are the sole responsibility of the authors and may not necessarily represent the official views of the Bill & Melinda Gates Foundation or other agencies that may have supported the primary data studies used in the present study.

Author Contributions

Conceptualization: Hans-Georg Müller.

Data curation: Michael S. Kramer, Seungmi Yang.

Formal analysis: Kyunghee Han.

Methodology: Kyunghee Han.

Project administration: Hans-Georg Müller.

Software: Kyunghee Han, Pantelis Z. Hadjipantelis.

Supervision: Hans-Georg Müller.

Validation: Kyunghee Han, Pantelis Z. Hadjipantelis.

Visualization: Kyunghee Han.

Writing – original draft: Kyunghee Han.

Writing – review & editing: Pantelis Z. Hadjipantelis, Jane-Ling Wang, Michael S. Kramer, Seungmi Yang, Richard M. Martin, Hans-Georg Müller.

References

1. Shenkin SD, Starr JM, Deary IJ. Birth weight and cognitive ability in childhood: A systematic review. *Psychological Bulletin*. 2004; 130(6):989. <https://doi.org/10.1037/0033-2909.130.6.989> PMID: 15535745
2. Morley R, Fewtrell MS, Abbott RA, Stephenson T, MacFadyen U, Lucas A. Neurodevelopment in children born small for gestational age: A randomized trial of nutrient-enriched versus standard formula and comparison with a reference breastfed group. *Pediatrics*. 2004; 113(3):515–521. <https://doi.org/10.1542/peds.113.3.515> PMID: 14993543
3. Räikkönen K, Forsén T, Henriksson M, Kajantie E, Heinonen K, Pesonen AK, et al. Growth Trajectories and Intellectual Abilities in Young Adulthood: The Helsinki Birth Cohort Study. *American Journal of Epidemiology*. 2009; 170(4):447–455. <https://doi.org/10.1093/aje/kwp132> PMID: 19528290

4. Yang S, Tilling K, Martin R, Davies N, Ben-Shlomo Y, Kramer MS. Pre-natal and post-natal growth trajectories and childhood cognitive ability and mental health. *International Journal of Epidemiology*. 2011; 40(5):1215–1226. <https://doi.org/10.1093/ije/dyr094> PMID: 21764769
5. Sudfeld CR, McCoy DC, Danaei G, Fink G, Ezzati M, Andrews KG, et al. Linear growth and child development in low-and middle-income countries: a meta-analysis. *Pediatrics*. 2015; 135(5):e1266–e1275. <https://doi.org/10.1542/peds.2014-3111> PMID: 25847806
6. Crookston BT, Schott W, Cueto S, Dearden KA, Engle P, Georgiadis A, et al. Postinfancy growth, schooling, and cognitive achievement: Young Lives. *The American Journal of Clinical Nutrition*. 2013; 98(6):1555–1563. <https://doi.org/10.3945/ajcn.113.067561> PMID: 24067665
7. Lundeen EA, Behrman JR, Crookston BT, Dearden KA, Engle P, Georgiadis A, et al. Growth faltering and recovery in children aged 1-8 years in four low- and middle-income countries: Young Lives. *Public Health Nutrition*. 2014; 17(9):2131–2137. <https://doi.org/10.1017/S1368980013003017> PMID: 24477079
8. Reid BM, Miller BS, Dorn LD, Desjardins C, Donzella B, Gunnar M. Early growth faltering in post-institutionalized youth and later anthropometric and pubertal development. *Pediatric Research*. 2017; 82(2):278–284. <https://doi.org/10.1038/pr.2017.35> PMID: 28170387
9. Keusch GT, Rosenberg IH, Denno DM, Duggan C, Guerrant RL, Lavery JV, et al. Implications of acquired environmental enteric dysfunction for growth and stunting in infants and children living in low-and middle-income countries. *Food and Nutrition Bulletin*. 2013; 34(3):357–364. <https://doi.org/10.1177/156482651303400308> PMID: 24167916
10. Gale CR, O'Callaghan FJ, Bredow M, Martyn CN, Avon Longitudinal Study of Parents and Children Study Team. The influence of head growth in fetal life, infancy, and childhood on intelligence at the ages of 4 and 8 years. *Pediatrics*. 2006; 118(4):1486–1492. <https://doi.org/10.1542/peds.2005-2629> PMID: 17015539
11. Lingam R, Hunt L, Golding J, Jongmans M, Emond A. Prevalence of developmental coordination disorder using the DSM-IV at 7 years of age: A UK population-based study. *Pediatrics*. 2009; 123(4):e693–e700. <https://doi.org/10.1542/peds.2008-1770> PMID: 19336359
12. Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, et al. Nurturing care: Promoting early childhood development. *The Lancet*. 2017; 389(10064):7–13. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3)
13. Hadjipantelis PZ, Han K, Wang JL, Yang S, Martin RM, Kramer MS, et al. Associating Growth in Infancy and Cognitive Performance in Early Childhood: A functional data analysis approach. *ArXiv e-prints*. 2018; 1808.01384.
14. Hall P, Müller HG, Wang JL. Properties of principal component methods for functional and longitudinal data analysis. *Annals of Statistics*. 2006; 34(3):1493–1517. <https://doi.org/10.1214/009053606000000272>
15. Wang JL, Chiou J, Müller HG. Functional Data Analysis. *Annual Review of Statistics and Its Application*. 2016; 3:257–295. <https://doi.org/10.1146/annurev-statistics-041715-033624>
16. Wei Y. An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts. *Journal of the American Statistical Association*. 2008; 103(481):397–409. <https://doi.org/10.1198/016214507000001472>
17. Zhang W, Wei Y. Regression based principal component analysis for sparse functional data with applications to screening growth paths. *The Annals of Applied Statistics*. 2015; 9(2):597–620. <https://doi.org/10.1214/15-AOAS811>
18. Park J, Ahn J. Clustering multivariate functional data with phase variation. *Biometrics*. 2017; 73(1):324–333. <https://doi.org/10.1111/biom.12546> PMID: 27218696
19. Cutler A, Breiman L. Archetypal analysis. *Technometrics*. 1994; 36(4):338–347. <https://doi.org/10.1080/00401706.1994.10485840>
20. Vinué G, Epifanio I, Alemany S. Archetypoids: A new approach to define representative archetypal data. *Computational Statistics & Data Analysis*. 2015; 87:102–2015. <https://doi.org/10.1016/j.csda.2015.01.018>
21. Epifanio I. Functional archetype and archetypoid analysis. *Computational Statistics & Data Analysis*. 2016; 104:24–34. <https://doi.org/10.1016/j.csda.2016.06.007>
22. Kramer MS, Chalmers B, Hodnett ED, Sevkovskaya Z, Dzikovich I, Shapiro S, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *Journal of the American Medical Association*. 2001; 285(4):413–420. <https://doi.org/10.1001/jama.285.4.413> PMID: 11242425
23. Kramer MS, Kakuma R. The optimal duration of exclusive breastfeeding. In: Pickering LK, Morrow AL, Ruiz-Palacios GM, Schanler R, editors. *Protecting Infants through Human Milk*. Springer; 2004. p. 63–77.

24. Tong S, McMichael AJ, Baghurst PA. Interactions between environmental lead exposure and sociodemographic factors on cognitive development. *Archives of Environmental Health*. 2000; 55:330–335. <https://doi.org/10.1080/00039890009604025> PMID: 11063408
25. Walker SP, Wachs TD, Gardner JM, Lozoff B, Wasserman GA, Pollitt E, et al. Child development: Risk factors for adverse outcomes in developing countries. *The Lancet*. 2007; 369(9556):145–157. [https://doi.org/10.1016/S0140-6736\(07\)60076-2](https://doi.org/10.1016/S0140-6736(07)60076-2)
26. Nisbett RE, Aronson J, Blair C, Dickens W, Flynn J, Halpern DF, et al. Intelligence: New findings and theoretical developments. *American Psychologist*. 2012; 67(2):130. <https://doi.org/10.1037/a0026699> PMID: 22233090
27. Smithers LG, Lynch JW, Yang S, Dahhou M, Kramer MS. Impact of neonatal growth on IQ and behavior at early school age. *Pediatrics*. 2013; 132(1):e53–e60. <https://doi.org/10.1542/peds.2012-3497> PMID: 23776123
28. Karhunen K. Zur Spektraltheorie Stochastischer Prozesse. *Annales Academiae Scientiarum Fennicae Series A I, Mathematica*. 1946; 7.
29. Loève M. Fonctions aléatoires à décomposition orthogonale exponentielle. *La Revue Scientifique*. 1946; 84:159–162.
30. Ramsay J, Silverman BW. *Functional data analysis*. 2nd ed. Springer-Verlag, New York; 2005.
31. Dauxois J, Pousse A, Romain Y. Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *Journal of Multivariate Analysis*. 1982; 12:136–154. [https://doi.org/10.1016/0047-259X\(82\)90088-4](https://doi.org/10.1016/0047-259X(82)90088-4)
32. Besse P, Ramsay JO. Principal components analysis of sampled functions. *Psychometrika*. 1986; 51:285–311. <https://doi.org/10.1007/BF02293986>
33. Silverman BW. Smoothed functional principal components analysis by choice of norm. *Annals of Statistics*. 1996; 24:1–24. <https://doi.org/10.1214/aos/1033066196>
34. Boente G, Fraiman R. Kernel-based functional principal components. *Statistics & Probability Letters*. 2000; 48:335–345. [https://doi.org/10.1016/S0167-7152\(00\)00014-6](https://doi.org/10.1016/S0167-7152(00)00014-6)
35. Hall P, Hosseini-Nasab M. On properties of functional principal components analysis. *Journal of the Royal Statistical Society: Series B*. 2006; 68:109–126. <https://doi.org/10.1111/j.1467-9868.2005.00535.x>
36. Li Y, Wang N, Carroll RJ. Selecting the number of principal components in functional data. *Journal of the American Statistical Association*. 2013; 108:1284–1291. <https://doi.org/10.1080/01621459.2013.788980>
37. James G, Hastie T, Sugar C. Principal component models for sparse functional data. *Biometrika*. 2000; 87:587–602. <https://doi.org/10.1093/biomet/87.3.587>
38. Rice JA, Wu CO. Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrika*. 2001; 57:253–259. <https://doi.org/10.1111/j.0006-341X.2001.00253.x>
39. Müller HG. Functional modeling and classification of longitudinal data. *Scandinavian Journal of Statistics*. 2005; 32(2):223–240. <https://doi.org/10.1111/j.1467-9469.2005.00429.x>
40. Yao F, Lee T. Penalized spline models for functional principal component analysis. *Journal of the Royal Statistical Society: Series B*. 2006; 68:3–25. <https://doi.org/10.1111/j.1467-9868.2005.00530.x>
41. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*. 2005; 100(470):577–590. <https://doi.org/10.1198/016214504000001745>
42. Yao F, Müller HG, Wang JL. Functional linear regression analysis for longitudinal data. *Annals of Statistics*. 2005; 33(6):2873–2903. <https://doi.org/10.1214/009053605000000660>
43. Müller HG. Functional modeling of longitudinal data. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. *Longitudinal Data Analysis*. CRC Press; 2009. p. 233–252.
44. Jones MC, Rice JA. Displaying the important features of large collections of similar curves. *The American Statistician*. 1992; 46(2):140–145. <https://doi.org/10.2307/2684184>
45. Leng X, Müller HG. Classification using functional data analysis for temporal gene expression data. *Bioinformatics*. 2006; 22(1):68–76. <https://doi.org/10.1093/bioinformatics/bti742> PMID: 16257986
46. Rousseeuw PJ, Ruts I, Tukey JW. The bagplot: A bivariate boxplot. *The American Statistician*. 1999; 53(4):382–387. <https://doi.org/10.2307/2686061>
47. Tukey JW. Mathematics and the picturing of data. In: James RD, editor. *Proceedings of the International Congress of Mathematicians*. vol. 2. Canadian Mathematical Society; 1975. p. 523–531.
48. Zuo Y. Projection-based depth functions and associated medians. *Annals of Statistics*. 2003; 31:1460–1490. <https://doi.org/10.1214/aos/1065705115>
49. Agostinelli C, Romanazzi M. Local depth. *Journal of Statistical Planning and Inference*. 2011; 141(2):817–830. <https://doi.org/10.1016/j.jspi.2010.08.001>

50. Hyndman RJ. Computing and graphing highest density regions. *The American Statistician*. 1996; 50(2):241–250. <https://doi.org/10.2307/2684423>
51. Scott DW. *Multivariate density estimation: Theory, practice and visualization*. 2nd ed. Wiley, New York; 2015.
52. Hyndman RJ, Shang HL. Rainbow plots, bagplots and boxplots for functional data. *Journal of Computational and Graphical Statistics*. 2010; 19:29–45. <https://doi.org/10.1198/jcgs.2009.08158>
53. Sun Y, Genton MG. Functional boxplots. *Journal of Computational and Graphical Statistics*. 2011; 20(2):316–334. <https://doi.org/10.1198/jcgs.2011.09224>
54. Hubert M, Rousseeuw PJ, Segaert P. Multivariate functional outlier detection. *Statistical Methods & Applications*. 2015; 24(2):177–202. <https://doi.org/10.1007/s10260-015-0297-8>
55. Narisetty NN, Nair VN. Extremal depth for functional data and applications. *Journal of the American Statistical Association*. 2016; 111(516):1705–1714. <https://doi.org/10.1080/01621459.2015.1110033>
56. Zhou L, Huang JZ, Carroll RJ. Joint modeling of paired sparse functional data using principal components. *Biometrika*. 2008; 95(3):601–619. <https://doi.org/10.1093/biomet/asn035> PMID: 19396364
57. Chiou JM, Müller HG. Linear manifold modeling of multivariate functional data. *Journal of the Royal Statistical Society: Series B*. 2011; 76(3):605–626. <https://doi.org/10.1111/rssb.12038>
58. Chiou JM, Chen YT, Yang YF. Multivariate functional principal component analysis: A normalization approach. *Statistica Sinica*. 2014; 24:1571–1596.
59. Happ C, Greven S. Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*. 2018; 113(522):649–659. <https://doi.org/10.1080/01621459.2016.1273115>
60. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995; p. 289–300.
61. Tukey JW. Comparing individual means in the analysis of variance. *Biometrika*. 1949; 5(2):99–114. <https://doi.org/10.2307/3001913>
62. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*. 1952; 47(260):583–621. <https://doi.org/10.1080/01621459.1952.10483441>
63. `fdapace`. R Package: Functional Data Analysis and Empirical Dynamics; version 0.4.0. Available from: <https://CRAN.R-project.org/package=fdapace>.
64. Emond AM, Blair PS, Emmett PM, Drewett RF. Weight faltering in infancy and IQ levels at 8 years in the Avon Longitudinal Study of Parents and Children. *Pediatrics*. 2007; 120(4):e1051–e1058. <https://doi.org/10.1542/peds.2006-2295> PMID: 17908725
65. Berrendero JR, Justel A, Svarc M. Principal components for multivariate functional data. *Computational Statistics & Data Analysis*. 2011; 55(9):2619–2634. <https://doi.org/10.1016/j.csda.2011.03.011>
66. Górecki T, Krzyśko M, Waszak Ł, Wołyński W. Selected statistical methods of data analysis for multivariate functional data. *Statistical Papers*. 2018; 59(1):153–182. <https://doi.org/10.1007/s00362-016-0757-8>
67. Cuevas A, Febrero M, Fraiman R. An anova test for functional data. *Computational Statistics & Data Analysis*. 2004; 47(1):111–122. <https://doi.org/10.1016/j.csda.2003.10.021>
68. Górecki T, Smaga Ł. Multivariate analysis of variance for functional data. *Journal of Applied Statistics*. 2017; 44(12):2172–2189. <https://doi.org/10.1080/02664763.2016.1247791>
69. Zhang JT. *Analysis of variance for functional data*. Chapman and Hall/CRC; 2013.
70. Sawant P, Billor N, Shin H. Functional outlier detection with robust functional principal component analysis. *Computational Statistics*. 2012; 27(1):83–102. <https://doi.org/10.1007/s00180-011-0239-3>
71. Febrero M, Galeano P, González-Manteiga W. Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels. *Environmetrics: The official journal of the International Environmetrics Society*. 2008; 19(4):331–345. <https://doi.org/10.1002/env.878>