Original Study

Precision Nutrition

OPEN

Defining longitudinal trajectory of body mass index percentile and predicting childhood obesity: methodologies and findings in the Boston Birth Cohort

Wanyu Huang[1], Anat Yaskolka Meir[2], Bolanle Olapeju[3], Guoying Wang[4], Xiumei Hong[4], Maya Venkataramani[5,6], Tina L. Cheng[7], Tak Igusa[1], Liming Liang[2,8,], Xiaobin Wang[4,5,*]*

[1]Department of Civil and Systems Engineering, Johns Hopkins University Whiting School of Engineering, Baltimore, MD, USA, [2]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [3]School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA, [4]Center on Early Life Origins of Disease, Department of Population, Family and Reproductive Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA, [5]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA, [6]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, [7]Cincinnati Children's Hospital Medical Center and Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA, [8]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

**Corresponding author. Address: Liming Liang, Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, USA. E-mail address: lliang@hsph.harvard.edu (L. Liang); Xiaobin Wang, Center on the Early Life Origins of Disease, Department of Population, Family and Reproductive Health, Johns Hopkins University Bloomberg School of Public Health, 615N. Wolfe Street, E4132, Baltimore, MD 21205, USA. E-mail address: xwang82@jhu.edu (X. Wang).*

Abstract

Background: Overweight or obesity (OWO) in school-age childhood tends to persist into adulthood. This study aims to address a critical need for early identification of children at high risk of developing OWO by defining and analyzing longitudinal trajectories of body mass index percentile (BMIPCT) during early developmental windows.

Methods: We included 3029 children from the Boston Birth Cohort (BBC) with repeated BMI measurements from birth to age 18 years. We applied locally weighted scatterplot smoothing with a time-limit scheme and predefined rules for imputation of missing data. We then used time-series *K*-means cluster analysis and latent class growth analysis to define longitudinal trajectories of BMIPCT from infancy up to age 18 years. Then, we investigated early life determinants of the BMI trajectories. Finally, we compared whether using early BMIPCT trajectories performs better than BMIPCT at a given age for predicting future risk of OWO.

Results: After imputation, the percentage of missing data ratio decreased from 36.0% to 10.1%. We identified four BMIPCT longitudinal trajectories: early onset OWO; late onset OWO; normal stable; and low stable. Maternal OWO, smoking, and preterm birth were identified as important determinants of the two OWO trajectories. Our predictive models showed that BMIPCT trajectories in early childhood (birth to age 1 or 2 years) were more predictive of childhood OWO (age 5–10 years) than a single BMIPCT at age 1 or 2 years.

Conclusions: Using longitudinal BMIPCT data from birth to age 18 years, this study identified distinct BMIPCT trajectories, examined early life determinants of these trajectories, and demonstrated their advantages in predicting childhood risk of OWO over BMIPCT at a single time point.

Keywords: Child health, Body mass index, Children BMI trajectory, Overweight or obesity, Data imputation, Longitudinal birth cohort, Multinomial logistic regression

Introduction

Overweight or obesity (OWO), defined as a body mass index (BMI) in the 85th percentile or greater for age and sex,[1] substantially raises the risk of morbidity and mortality as a result of serious related health complications such as type 2 diabetes and coronary heart disease.[2,3] Given its link with adult OWO,[4] childhood OWO is now a major health concern worldwide.[5,6] The overall prevalence of obesity for children and adolescents aged 2 to 19 years in the US was 19.7% in 2017 to 2020, affecting over 14 million children.[7] Hence, the study of childhood OWO and associated antecedent factors has become a national health priority.

Most epidemiologic studies[5,8–11] to date of childhood OWO have focused on childhood BMI at one time point only, ignoring

Liming Liang and Xiaobin Wang contributed equally to the work.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Clinical Trial Registry: Boston Birth Cohort Study, NCT03228875. https://clinicaltrials.gov/ct2/show/NCT03228875Supplemental Digital Content is available for this article.

Precis Nutr 2023;2(2):e00037.

Received: 6 February 2023, Revised: 11 March 2023, Accepted: 26 March 2023

doi: 10.1097/PN9.0000000000000037

the dynamic changes of child development and thus prone to biased or incomplete conclusions.[12,13] A more rigorous dynamic analysis of BMI growth patterns during childhood is possible when using statistical methods for long-span longitudinal data.[14] Recent examples of such methods include time series *K*-means clustering,[15–17] latent class analysis,[18] and latent class growth analysis (LCGA).[19] While a few studies[12,20–24] have prospectively examined childhood BMI trajectories, all of them were limited to the exploration of a specific age range, for example, age 0 to 2 years,[21] or 0 to 12 years.[12] Another key concern is the way BMI records are used for analysis in these studies. As children's weight and length/height data used to determine BMI are typically obtained from public records[25] or cohort studies,[21] these longitudinal data points are likely to have missing values or irregular time intervals. In such instances, researchers either choose to limit their longitudinal analyses to complete cases[26] without missing data, which reduces statistical power,[23] or apply imputation methods.[21,24] Furthermore, there has been little quantitative analysis of BMI longitudinal data in high risk but understudied subgroups in the US, such as minority and low-income populations.[15,27] Hence, there is a need for statistical methods that can provide a deeper and more rigorous analysis of BMI trajectories among high-risk populations to mitigate the public health burden of childhood OWO in the US.

This study aims to address a critical need for early identification of children at high risk of developing OWO by defining and analyzing longitudinal trajectories of BMI percentile (BMIPCT) during early developmental windows. We also explored associated factors of these BMIPCT growth trajectories and assessed their ability to predict OWO risks at later life stages. A notable strength of this study was to the use of our rigorous statistical methods to analyze longitudinal data from the Boston Birth Cohort (BBC),[28] a US urban low-income, predominantly Black and Hispanic minority population, who are disproportionally affected by OWO.

Methods

This study included 3029 children of BBC who were enrolled at birth and followed prospectively up to age 21 years, with repeated measurements of body weight and length/height. Supplementary Figure S1, http://links.lww.com/PN9/A24, depicts the study population flowchart including inclusion and exclusion criteria. The BBC has been previously described in detail.[28,29] Briefly, research staff at Boston Medical Center (BMC) recruited mothers 24 to 72 hours after delivery and obtained written informed consent for the long-term longitudinal study. It initially enrolled 8509 mother–infant dyads from 1998 to 2016. As of July 2018, 3029 children had postnatal follow-up. Their height and weight were measured during well-child visits based on standard clinical procedures as part of the pediatric primary care schedule and were recorded in their electronic medical records (EMR) at BMC.[30]

Before age 2 years, BMIPCT were calculated based on World Health Organization reference values.[31] In subsequent years, BMIPCT were defined based on US national reference values from the Centers for Disease Control and Prevention (CDC).[32] Mother–child dyads in the BBC visited at different frequencies, for example, for babies and toddlers younger than age 2 years, well-child visits were much more frequent (multiple times per year) compared with annual well-child care visits after age 2 years. According to primary care visit schedules and knowledge of child growth, our study defined 32 time intervals from birth to age 18 years as follows: 12 monthly time intervals from birth to age 1 year; 4 quarterly time intervals from age 1 to 2 years; and 16 annual time intervals afterward. BMIPCT-related values for each time interval were defined by the average values in the EMR during the corresponding timeframe.

Maternal variables, including age at time of delivery, race and ethnicity, education level, pre-pregnancy height and weight, smoking status, parity, and mode of delivery, were obtained from medical records and a maternal questionnaire interview. Maternal pre-pregnancy BMI was calculated from pre-pregnancy height and weight and further categorized into non-OWO (BMI < 25) and OWO (BMI ≥ 25). Breastfeeding was assessed during the first 2 years of follow-up[33] (any *vs.* no breastfeeding). The child's sex, birthweight, and preterm birth status were obtained from the EMR.

Statistical analyses

Missing data imputation using locally weighted scatterplot smoothing with a time-limit aware scheme

We applied locally weighted scatterplot smoothing with a time-limit aware scheme (LOWESS-TLS) to impute the data. While LOWESS imputation has been used in geographical research, its use in BMI longitudinal studies is novel; moreover, our proposed combination of LOWESS with TLS is a novel method. Hence, we describe this two-step imputation method in detail below.

TLS is the first step in the imputation procedure. In most longitudinal BMI studies, imputation is performed using all available data or data within a fixed-width time window centered on the time of the missing data. Herein, we use TLS to account for the age-varying dynamics of child growth. For missing monthly data from birth to age 1 year, imputation was performed using data no more than 4 months earlier or later; for missing quarterly data between age 1 and 2 years, data no more than 1 year earlier or later is used; for missing annual data between age 3 and 18 years, data no more than 2 years earlier or later is used. These variable-width time windows are schematically presented in Supplementary Figure S2, http://links.lww.com/PN9/A24, with an illustrative example.

The second imputation step begins with the set of available data points within the TLS. When this set includes data points with ages both above and below that of the missing data, then LOWESS regression on the set of observations was used to predict the missing value.[34] For cases when only the last or next observations are available, last-observation-carried-forward (LOCF) or next observation carried backwards (NOCB)[35] imputation was conducted. Finally, if there were no known BMIPCT measurements in the TLS time window, the missing data were not imputed and any subsequent analysis that requires this particular data point would not include this participant. Once all of the LOWESS-TLS imputations were completed for each participant, we linearly interpolated the imputed dataset to $18 \times 12 = 216$ monthly values to simplify the subsequent trajectory analysis.

BMIPCT trajectory identification and analysis

The LOWESS-TLS imputed, linearly interpolated data was used to create 18 BMIPCT datasets as follows: dataset *n* (for

n = 1, 2, …, 18) contains data only up to age n years and does not contain any missing data that could not be imputed using our TLS criteria, as explained above. For each of these BMIPCT datasets, we applied time-series K-means clustering[36] to obtain trajectories. To determine the number of trajectories to retain for subsequent analysis, we considered a range of possible numbers of trajectories (ie, the number of clusters k = 2, …, 10 in K-means clustering), computed the Bayesian Information Criterion (BIC),[11,37,38] and selected the number k corresponding to the highest BIC. Euclidean distance was chosen for K-means clustering since our time series data were of equal durations and equally spaced time intervals. For comparison, we applied LCGA[39] as an alternative method and set the number of latent trajectory classes in LCGA the same as k for consistency. To allow for sex-specific trajectories and associations, all analyses were performed on the overall study population as well as on boys and girls, separately.

To examine associations between BMIPCT trajectory types and maternal factors, child sex, preterm and low birthweight status, we applied a two-sided t test for continuous factors and the χ^2 test for categorical factors. We also conducted multinomial logistic regression, taking the identified normal trajectory as the reference category, and reported crude and adjusted odds ratios (ORs and aORs).

Prediction of OWO risk in middle childhood

After an investigation of the trajectory classes, we explored how trajectory information could be used to predict subsequent OWO risk. To accomplish this, we developed a preliminary prediction tool using early-age information (from birth to age 1 or 2 years) to inform parents about their child's risk of OWO at a critical developmental time window (from age 5 to 10 years),[40,41] with and without further consideration of prenatal and perinatal covariates. Binary logistic regression was used to predict middle-childhood OWO status. To assess if a child's early growth patterns (as represented by BMIPCT trajectories from birth to age n_1 years, n_1 = 1,2) could better predict middle-childhood OWO status at age n_2 years (n_2 = 5,6,…,10) compared with conventional predictors or static BMI categories, we chose five different sets of predictors and collated their predictive performances on middle childhood OWO: (a) baseline maternal and infant factors, including maternal OWO status, race and ethnicity, education level, parity, smoking status, delivery type, preterm status, low birthweight, and breastfeeding type; (b) static BMIPCT measurement at age n_1 years, categorized into four levels: <5th, 5th to 85th, 85th to 95th, or ≥95th; (c) BMIPCT trajectory types from birth to age n_1 years; (d) the set of all predictors in sets (a) and (b); and (e) the set of all predictors in sets (a) and (c). Using five-fold cross-validation (CV), we reported the area under the receiver operating characteristics curve (AUC) as the performance metric of each set of predictors.

All statistical analyses were performed using Python (version 3.9.13; Python Software Foundation) and R (version 3.6.3; R Development Core Team, Vienna, Austria).

Results

Data imputation

The subset of the BBC longitudinal data that is used in this OWO study was closed on July 31, 2018. Since the birth dates ranged from 1998 to 2016, most of the children were below 18 years old at the time the dataset was closed, resulting in administrative right censoring. Of the 96,928 possible time steps (3029 participants multiplied by 32 time steps per participant), 28,665 (29.6%) were administratively right censored (where the percentage corresponds to all possible time steps). Of the remaining 68,263 time steps, there were 34,894 (36.0% of all possible data) missing data. There were two causes of missing data in our data set: no well-child visit scheduled at one of the 32 time points and mother–child dyads missing one or more well-child visits. The missing data are graphically shown in the data availability matrix Figure 1A in which the columns represent the 32 time steps (monthly from birth to age 1 year, quarterly from age 1 to 2 years, and annual from age 2 to 18 years) and the rows represent the 3029 dyads. In each cell of this rectangular array, green indicates that BMIPCT data were available, purple indicates missing data, and yellow corresponds to the administratively censored data.

By applying LOWESS-TLS, 25,135 (25.9% of all possible data) of the non-administrative missing data and 6638 (6.9%) of the administrative data were imputed, and only 10.1% were missing. Figure 1B shows the updated data availability matrix after applying LOWESS-TLS. Supplementary Figure S3A, http://links.lww.com/PN9/A24, shows a box plot of the comparison of the mean and median values between the raw data and LOWESS-TLS imputed data, indicating that the statistics remained comparable and consistent. As such, in the subsequent analyses, we exploited the LOWESS-TLS imputed BMIPCT dataset to increase statistical power while retaining similar statistical properties as the raw data.

The number of study samples in each of the 18 BMIPCT datasets, where dataset n corresponds to data points from birth to age n years, is shown in orange in Supplementary Figure S4, http://links.lww.com/PN9/A24. It can be seen that as n increases, the number of samples decreases. This is because of the requirement that a child's data would be retained in the dataset only if all of the missing data could be imputed under the TLS criteria. Supplementary Figure S4, http://links.lww.com/PN9/A24, also shows the number of study samples when the 18 BMIPCT datasets are constructed using other procedures, including complete cases (no imputation) and data aggregated to 18 yearly time points. We can see that the use of LOWESS-TLS imputation with 32 time points provides far more samples than the complete cases (eg, 1158 *vs.* 0 complete cases for birth to age 10 years), even when aggregated to 18 yearly time points (eg, 1323 *vs.* 221 complete cases for birth to age 10 years). While aggregating the LOWESS-TLS imputed data from 32 to 18 time points provides slightly more samples (eg, 2154 *vs.* 2607 at age 1 year, where the difference is highest), it is felt that the shorter time intervals used in the first 2 years of the 32 time points are needed to capture the details of early child growth dynamics.

As age 10 years represents middle childhood[42] and has shown high predictability of OWO in adolescence and adulthood,[6,43,44] our primary focus is on the BMIPCT dataset from birth to age 10 years. As indicated in Supplementary Figure S4, http://links.lww.com/PN9/A24, there are 1158 children in the LOWESS_TLS imputed dataset n = 10, where all included

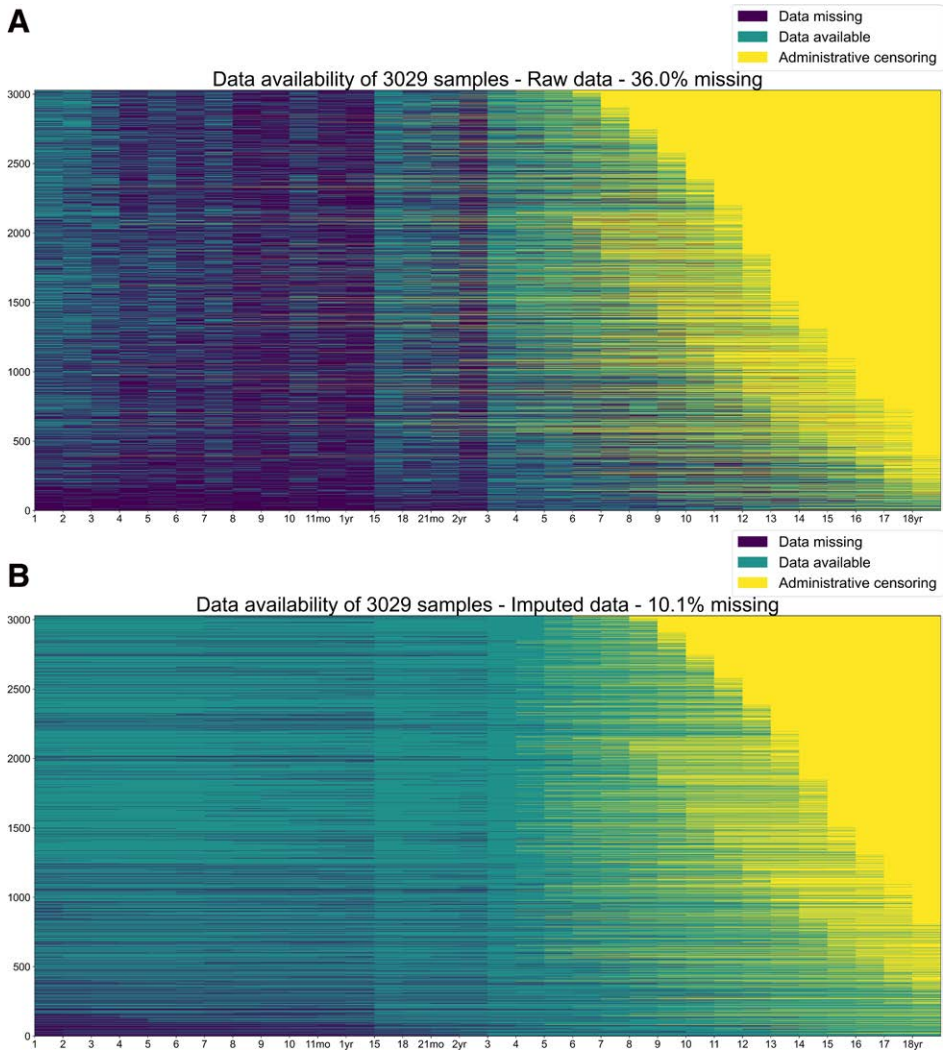

Figure 1: Data availability before imputation (top panel) and after imputation* (bottom panel) for all samples. X-axis is the time (monthly data from birth to age 1 year, quarterly data from age 1 to 2 years, and annual data from age 2 to 18 years), and each row of the Y-axis represents one child. In each column, green indicates BMIPCT data is available, purple means BMIPCT data is missing, and yellow represents administratively censored data, which are the points beyond participants' ages. *Time-limit aware scheme in data imputation as illustrated in Supplementary Figure S2, http://links.lww.com/PN9/A24: for missing monthly data from birth to age 1 year, imputation was allowed using data no more than 4 months earlier/later; for missing quarterly data between age 1 and 2 years, collected data no more than 1 year apart could be used; for subsequent missing annual data, imputation could be carried out using data no more than 2 years apart. BMIPCT, body mass index percentile.

children have no missing data from birth to age 10 years after imputation.

Population characteristics

Table 1 shows the population characteristics, stratified by child sex, of the imputed dataset from birth to age 10 years. Of the 1158 study children, 586 (50.6%) were boys and over half (52.8%) of the mothers were OWO during pre-pregnancy. As shown, over 70% of the mothers self-identified as being of Black race, and the mean (SD) maternal age at birth of the study children was 29.1 (6.81) years. Among the mothers, 64.6% did not reach a college degree before the birth of the enrolled child, and most (76.2%) never smoked. There were 341 (29.4%) children identified as OWO at age 10 years. In Supplementary Table S1, http://links.lww.com/PN9/A24, we also report the population characteristics of the imputed dataset from birth to age 2 years (N = 2124 study dyads), which is a sensitive and critical window for the development of later childhood OWO.[21]

BMIPCT growth trajectories and its associated factors

We conducted K-means clustering analysis on each of the 18 BMIPCT datasets (from birth to age n years where n = 1, 2, …, 18). To determine the number of trajectories (ie, number of clusters), we compared BIC results for n = age 10 years for k = 2, …, 10 clusters (Supplementary Table S2, http://links.lww.com/PN9/A24).[11,37,38] The results show that k = 4 trajectories corresponds to the highest BIC value, which is in agreement with prior childhood BMI trajectory analyses.[13,15] For consistency, four trajectories were used for all of the 18 BMIPCT datasets.

The four BMIPCT trajectories for birth to age 10 years are shown in Figure 2A with labels that indicate their distinct characteristics, as described in the following. Early onset OWO (43.7% of the children) exhibits high BMIPCT in early childhood that was maintained throughout childhood. Late onset OWO (20.4%) shows a rapid BMIPCT gain starting at age 3 years before attaining high BMIPCT in later childhood. Normal stable (16.7%) demonstrated a consistently normal BMIPCT,

Table 1

Population characteristics of mother–child pairs, who have complete BMIPCT data points from birth to age 10 years after imputation.

Characteristic	Overall sample (*N* = 1158)	Girls (*N* = 572)	Boys (*N* = 586)
Mothers with OWO (%)	611 (52.8)	297 (51.9)	314 (53.6)
Maternal age (mean [SD])	29.11 (6.81)	28.98 (6.76)	29.24 (6.85)
Race and ethnicity (%)			
Black	856 (73.9)	411 (71.9)	445 (75.9)
White	50 (4.3)	27 (4.7)	23 (3.9)
Hispanic	188 (16.2)	106 (18.5)	82 (14.0)
Other	64 (5.5)	28 (4.9)	36 (6.1)
Maternal education, ≥college (%)	410 (35.4)	203 (35.5)	207 (35.3)
Maternal smoking, continuous & quitter (%)	272 (23.8)	139 (24.7)	133 (22.9)
Parity, multiparous (%)	702 (60.6)	342 (59.8)	360 (61.4)
Mode of delivery, C-section (%)	733 (63.3)	372 (65.0)	361 (61.6)
Preterm birth (%)	273 (23.6)	132 (23.1)	141 (24.1)
Low birthweight (%)	266 (23.0)	144 (25.2)	122 (20.8)
Breastfeeding type, formula feeding only (%)	273 (24.5)	129 (23.5)	144 (25.4)
Children with OWO, defined by the BMIPCT at age 10 years (%)	341 (29.4)	169 (29.5)	172 (29.4)

BMIPCT, body mass index percentile; C-section, cesarean section; OWO, overweight or obesity.

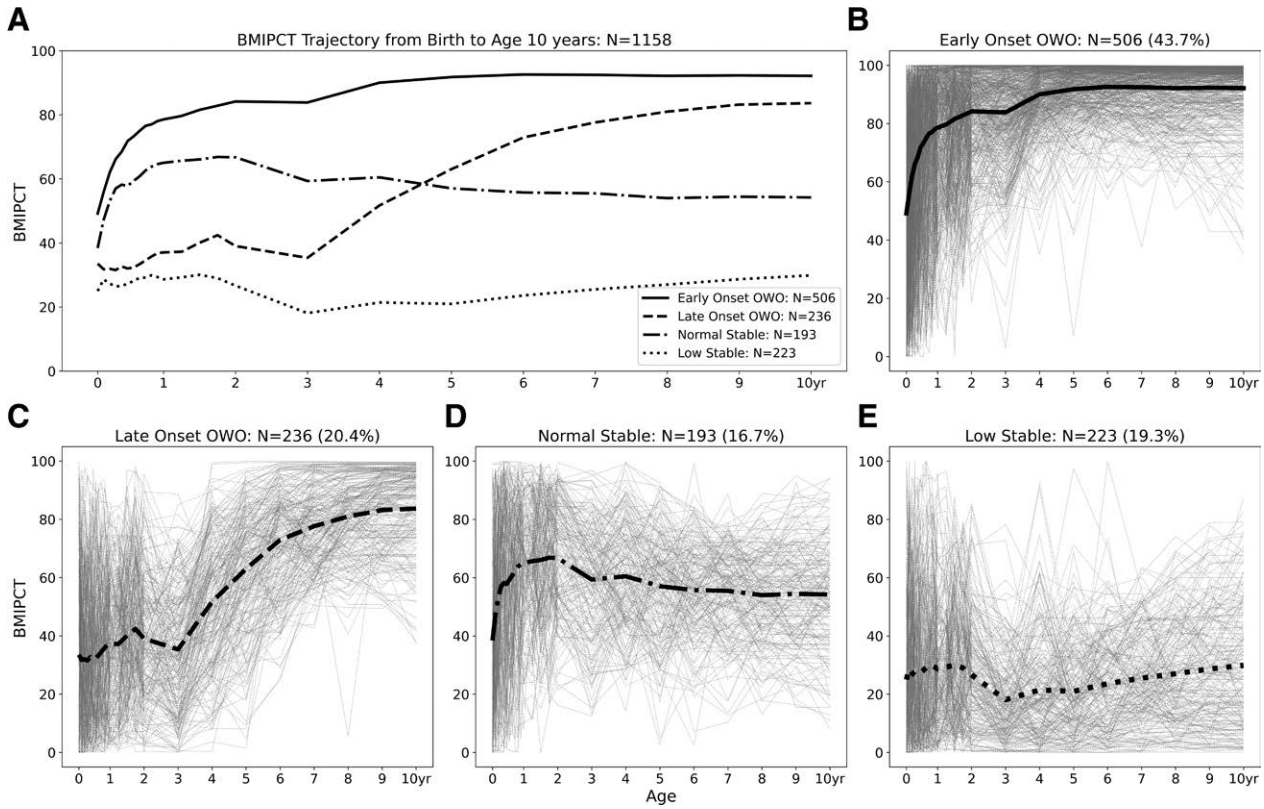

Figure 2: Summary of four BMI trajectory groups from birth to age 10 years. Thick black lines indicate the mean trajectories identified for early onset OWO (solid); late onset OWO (dashed); normal stable (dash-dotted); and low stable group (dotted). Actual trajectories for participants in each defined trajectory group are shown by the thin gray line in (B) to (E). BMI, body mass index; OWO, overweight or obesity.

while low stable (19.3%) maintained a stable but relatively lower BMIPCT, typically around the 20th percentile, within the time window.

With the clustering method, each child is assigned to a unique trajectory, resulting in a stratification of the 1158 children. The individual BMIPCT trajectories and centroids for each of the four trajectories are shown in Figure 2B–E. The population demographic statistics for the 1,158 study samples, stratified by growth trajectories, are shown in Supplementary Table S3, http://links.lww.com/PN9/A24. Using the chi-squared test, we found statistically significant association of maternal OWO, maternal age, race and ethnicity, C-section delivery mode, and low birthweight with child BMIPCT trajectories. We used multinomial logistic regression to explore these associations further. Table 2 shows the associations between all studied maternal and child factors and BMIPCT growth trajectory patterns from birth to age 10 years, reported by the ORs (aORs) and 95% confidence intervals (CIs), using the Normal Stable trajectory as the reference group. Results on boys and girls subgroups are also shown in Supplementary Table S4, http://links.lww.com/PN9/A24. Compared to children born to non-OWO mothers, children born

Table 2

Associations of maternal and child factors with child BMIPCT trajectory groups from birth to age 10 years based on multinomial logistic regression model results.

Characteristics	OR* (95% CI)	aOR† (95% CI)	Characteristics	OR (95% CI)	aOR (95% CI)
Maternal OWO (non-OWO [ref] *vs.* OWO mothers)			Parity (nulliparous [ref] *vs.* multiparous)		
Early onset OWO	2.07 (1.48–2.90)	1.68 (1.18–2.40)	Early onset OWO	1.13 (0.89–1.43)	0.91 (0.68–1.20)
Late Onset OWO	2.01 (1.37–2.96)	1.76 (1.17–2.65)	Late onset OWO	1.07 (0.81–1.41)	0.87 (0.63–1.21)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	0.79 (0.53–1.17)	0.81 (0.54–1.23)	Low stable	0.90 (0.69–1.19)	0.89 (0.64–1.24)
Maternal age (per 1-year increase)			Mode of delivery (vaginal [ref] *vs.* C-section)		
Early onset OWO	1.04 (1.01–1.06)	1.03 (1.00–1.06)	Early onset OWO	0.73 (0.52–1.03)	0.82 (0.57–1.19)
Late onset OWO	1.04 (1.01–1.07)	1.04 (1.01–1.08)	Late onset OWO	0.94 (0.63–1.41)	1.21 (0.79–1.87)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	1.00 (0.97–1.03)	1.01 (0.97–1.04)	Low stable	1.44 (0.95–2.20)	1.51 (0.96–2.37)
Race/ethnicity (Black [ref] *vs.* White)			Preterm birth (full term [ref] *vs.* preterm)		
Early onset OWO	0.68 (0.30–1.57)	0.56 (0.23–1.34)	Early onset OWO	1.11 (0.74–1.67)	1.68 (0.97–2.90)
Late onset OWO	1.28 (0.55–3.02)	1.15 (0.46–2.86)	Late onset OWO	1.50 (0.95–2.36)	1.57 (0.86–2.87)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	0.84 (0.32–2.18)	0.80 (0.30–2.15)	Low stable	1.39 (0.87–2.20)	1.43 (0.78–2.61)
Race/ethnicity (Black [ref] *vs.* Hispanic)			Low birth weight (non-LBW [ref] *vs.* LBW)		
Early onset OWO	0.98 (0.64–1.52)	1.05 (0.67–1.66)	Early Onset OWO	0.65 (0.43–0.96)	0.45 (0.26–0.77)
Late onset OWO	0.66 (0.39–1.13)	0.72 (0.41–1.27)	Late Onset OWO	1.18 (0.76–1.83)	0.97 (0.54–1.74)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	0.67 (0.39–1.16)	0.74 (0.42–1.30)	Low stable	1.33 (0.86–2.06)	1.12 (0.63–1.98)
Race/ethnicity (Black [ref] *vs.* others)			Breastfeeding type (any breast feeding [ref] *vs.* formula feeding only)		
Early onset OWO	0.48 (0.22–1.05)	0.49 (0.22–1.07)	Early onset OWO	1.08 (0.73–1.58)	1.10 (0.73–1.66)
Late onset OWO	0.83 (0.37–1.88)	0.76 (0.33–1.79)	Late onset OWO	0.96 (0.61–1.50)	1.05 (0.65–1.68)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	1.61 (0.77–3.36)	1.68 (0.78–3.60)	Low stable	0.74 (0.46–1.19)	0.73 (0.45–1.20)
Maternal education (≤high school [ref] *vs.* ≥college)			Child sex (female [ref] *vs.* male)		
Early onset OWO	0.94 (0.66–1.33)	0.87 (0.60–1.27)	Early onset OWO	0.85 (0.61–1.18)	0.82 (0.58–1.16)
Late onset OWO	0.74 (0.50–1.10)	0.77 (0.50–1.18)	Late onset OWO	0.90 (0.62–1.32)	0.94 (0.63–1.41)
Normal stable	1 (reference)	1 (reference)	Normal stable	1 (reference)	1 (reference)
Low stable	0.96 (0.64–1.44)	0.99 (0.63–1.54)	Low stable	0.82 (0.56–1.20)	0.84 (0.56–1.25)
Maternal smoking (never [ref] *vs.* continuous & quitter)					
Early onset OWO	1.08 (0.82–1.42)	1.15 (0.84–1.56)			
Late onset OWO	0.98 (0.71–1.35)	0.95 (0.66–1.35)			
Normal stable	1 (reference)	1 (reference)			
Low stable	1.05 (0.76–1.45)	1.18 (0.83–1.68)			

Significance level is 5%. As this is 95% CI, and bold value means it does not contain/cross the value 1, so the P value is strictly less than 0.05.

*Odds ratio, where only examined independent variable is included in the univariate multinomial logistic regression.

†Adjusted odds ratio, where all predictor variables are simultaneously included in the multivariate multinomial logistic regression model.

aOR, adjusted odds ratio; BMIPCT, body mass index percentile; C-section, cesarean section; CI, confidence interval; LBW, low birthweight; OR, odds ratio; OWO, overweight or obesity; SD, standard deviation.

to mothers with OWO have a 1.68 (95% CI: 1.18–2.40) times higher chance of becoming early onset OWO and a 1.76 (95% CI: 1.17–2.65) times higher chance of becoming late onset OWO, with the normal stable group as a reference. Besides, maternal age is positively linked with the odds of being early onset OWO (aOR: 1.03, 95% CI: 1.00–1.06) and late onset OWO (aOR: 1.04, 95% CI: 1.01–1.08) with normal stable serving as a reference. The probability of children with low birth weight to be classified as early onset OWO with respect to being normal stable was 0.45 times lower than children without low birth weight (95% CI: 0.25–0.77). Babies born preterm, particularly boys, have higher odds of rapidly increasing BMI (late onset OWO) (aOR: 3.39, 95% CI: 1.39–8.26) as opposed to maintaining a normal BMI in the first 10 years, when compared to boys born full term.

While our primary focus is on birth to age 10 years trajectories, we also report results for trajectories for other age ranges in the Supplement. Supplementary Figure S5, http://links.lww.com/PN9/A24, illustrates the four trajectories of each BMIPCT dataset (from birth to n years, n = 1, 2,..., 18), and the growth patterns of the four types identified in different age stages are similar, but they hold distinct clinical implications. For instance, children who exhibit late onset OWO before the age of 2 are likely to be classified as early onset OWO within the time window of 0 to 10 years. Supplementary Figure S6, http://links.lww.com/PN9/A24, is a Sankey diagram that shows the transition flows of BMIPCT trajectories throughout the three age points (birth to 1 year, to 2 years, to 10 years), for example, most children who were classified as early onset OWO or low stable at 1 or 2 years old remained in the same groups at 10 years old. Supplementary Table S5, http://links.lww.com/PN9/A24, shows descriptive statistics of 2124 samples stratified by the four trajectories from birth to age 2 years, and Supplementary Table S6, http://links.lww.com/PN9/A24 shows factors associated with these four trajectories. It can be seen that, for this 2-year time window, children of mothers with OWO were less likely to follow a low stable trajectory pattern in the first 2 years (aOR: 0.66, 95% CI: 0.50–0.87). Besides, maternal smoking is

associated with an increased risk of early onset OWO (aOR: 1.39, 95% CI: 1.12–1.73).

Predictions of OWO risk at age 5 to 10 years using early childhood BMIPCT trajectories

Next, we demonstrate how childhood BMIPCT trajectories from birth to age 1 or 2 years can be used to predict OWO risks at ages 5 through 10 years. We conduct a series of logistic regression of child OWO status on five sets of predictors as described in Methods section and use the AUC of five-fold CV to assess prediction performance. For the base model (a), we use only conventional clinical measurements as predictors (including maternal and infant factors collected immediately after birth). Next, we assess two crude models, with model (b) using static BMI categories as predictors and model (c) using BMIPCT trajectories as predictors. Finally, we repeat models (b) and (c) while adjusting for all of the covariates considered in model (a) and refer to these as adjusted models (d) and (e), respectively. As shown in Table 3, the AUCs of the base models range from 0.625 to 0.648, which are lower than predictive performances of all crude models and adjusted models. More importantly, our results indicated that child BMIPCT developmental trajectories from birth to age 1 or 2 years were more informative than static BMI categories at age 1 or 2 years for predicting OWO status at ages 5 to 10 years and that the improvement persisted when using adjusted models. For example, when predicting OWO status at age 6 years for the 1802 samples, the AUC of five-fold CV was 0.626 for the base model using factors (a), 0.664, and 0.714 for the crude models using predictors (b) and (c), respectively, and was 0.718 and 0.736 for the adjusted models (d) and (e). Similar improvements were observed in the subgroup analysis by child sex. The predictive performance for predicting OWO at ages 11 to 18 years are reported in Supplementary Table S7, http://links.lww.com/PN9/A24.

Sensitivity analysis

LCGA, a special type of growth mixture modeling, whereby the variance and covariance estimates for the growth factors within each class are assumed to be zero, is another commonly used approach to BMI trajectory analysis. The BMIPCT growth trajectories identified by LCGA, shown in Supplementary Figure S7, http://links.lww.com/PN9/A24, were comparable and consistent with our *K*-means results, as shown in the pairwise plot (Supplementary Figure S8, http://links.lww.com/PN9/A24). This finding further illustrates the robustness of the BMIPCT trajectories in our *K*-means model. We also compared the predictive performances of the BMIPCT trajectories from birth to age 1 or 2 years as identified by *K*-means clustering and LCGA for middle childhood OWO risks. The results showed close predictions, as shown in Supplementary Table S8, http://links.lww.com/PN9/A24.

We also repeated the trajectory analysis and the explorations of the association between early childhood BMIPCT trajectory and middle childhood OWO status on the imputed dataset using the simple average with the TLS. In other words, compared with the LOWESS-TLS, we replaced the step using the LOWESS curve with the average of the previous and subsequent observations within the TLS. Average-TLS imputed data produced larger discrepancies with the observed values' mean and standard deviation, compared with LOWESS-TLS. Supplementary Figure S3B, http://links.lww.com/PN9/A24, shows the box plot including the mean and median values between the raw data and simple average imputed data. Supplementary Table S9, http://links.lww.com/PN9/A24, shows the predictive power of the early childhood trajectories from the simple average imputed dataset on middle childhood OWO risks. Results for the simple average imputation methods are comparable to the results we obtained based on LOWESS-TLS.

Discussion

In the BBC cohort, consisting primarily of Black underrepresented population from urban and largely low-income communities, we identified four distinct BMIPCT developmental trajectories, including early onset OWO, late onset OWO, normal stable, and low stable, from birth to age *n* years ($n = 1, 2, \ldots, 18$) with a primary focus on birth to age 10 years. We found that trajectory patterns identified using *K*-means clustering are consistent with those from LCGA, which supports the robustness of the four types of BMIPCT trajectories. Aris *et al*[21] and Bichteler and Gershoff[25] also identified similar BMI trajectories to ours at two different age windows, birth to age 2 years and 2 years to 13 years, respectively, although they used different statistical methods and studied different populations. The trajectories from birth to age *n* years during childhood can be used by clinicians to more readily observe potential dynamic changes that can emerge during child development and identify rapid weight gain patterns among their patients throughout childhood.[13] It can also help identify subgroups of children who are at higher risk of developing obesity and target interventions to these groups. Additionally, understanding the predictors and factors associated with different BMI trajectories can help inform policies and programs aimed at reducing the prevalence of childhood obesity.

Our results indicate that maternal OWO is a strong and dominant risk factor for increased BMI, including early onset and late onset weight gain, in the first 10 years of childhood, which is consistent with previous findings.[27,45] Genetic, epigenetic, and shared environmental characteristics may contribute to this association.[46] In addition, boys born preterm were more likely to experience late onset OWO with respect to being normal stable, compared to boys born full term. This is in line with earlier studies suggesting that preterm infants often exhibit rapid "catch-up" growth during the initial years of life, and rapid weight gain during this period is linked to a higher likelihood of childhood obesity.[47] We also find maternal smoking is associated with an increased risk of early onset OWO for children before age 2 years, but we did not detect a significant association with BMIPCT trajectories at age 10 years. This lack of association could be attributed to the influence of other factors, such as physical activity[48] and socioeconomic status,[49] on child BMI development, which may intensify with age. Additionally, the effects of maternal smoking on BMI may be more pronounced in early childhood and fades with age.[50,51] Another possibility is that the sample size was not sufficient to detect a significant association between maternal smoking and BMI trajectories at age 10 years. Future studies should consider using larger sample sizes and longer follow-up periods, as well as taking into account other potential confounding variables, in order to fully understand the relationship between maternal smoking and BMI trajectories in children. The information on the identified predictors of BMIPCT trajectories could be useful in guiding clinicians in identifying children at higher risk for OWO during childhood. In addition, we used early childhood growth patterns to predict the risk of OWO in middle childhood. Our

Table 3

Prediction performance on OWO at different ages (5–10 years).

	Crude model*		Adjusted model†	
Five-fold CV AUC mean (SD)	**BMI category @*n* years**	**BMI trajectory @0 to *n* years**	**BMI category @*n* years**	**BMI trajectory @0 to *n* years**
Overall samples				
Y: OWO status @5 years (*N* = 1929) (base model‡-0.632 [0.041])				
n = 1	0.681 (0.008)	0.729 (0.007)	0.737 (0.021)	0.751 (0.012)
n = 2	0.739 (0.025)	0.782 (0.007)	0.781 (0.021)	0.799 (0.009)
Y: OWO status @6 years (*N* = 1802) (base model-0.626 [0.028])				
n = 1	0.664 (0.016)	0.714 (0.014)	0.718 (0.015)	0.736 (0.022)
n = 2	0.722 (0.027)	0.76 (0.013)	0.763 (0.03)	0.78 (0.015)
Y: OWO status @7 years (*N* = 1702) (base model-0.632 [0.026])				
n = 1	0.638 (0.017)	0.691 (0.028)	0.696 (0.012)	0.717 (0.03)
n = 2	0.703 (0.035)	0.736 (0.029)	0.749 (0.034)	0.758 (0.028)
Y: OWO status @8 years (*N* = 1546) (base model-0.629 (0.024))				
n = 1	0.627 (0.011)	0.676 (0.018)	0.683 (0.016)	0.702 (0.027)
n = 2	0.689 (0.021)	0.718 (0.012)	0.732 (0.032)	0.739 (0.021)
Y: OWO status @9 years (*N* = 1328) (base model-0.627 [0.037])				
n = 1	0.622 (0.024)	0.672 (0.029)	0.673 (0.034)	0.694 (0.027)
n = 2	0.671 (0.032)	0.71 (0.023)	0.714 (0.04)	0.727 (0.024)
Y: OWO status @10 years (*N* = 1158) (base model-0.648 [0.04])				
n = 1	0.613 (0.031)	0.661 (0.042)	0.683 (0.046)	0.702 (0.036)
n = 2	0.661 (0.021)	0.696 (0.034)	0.718 (0.036)	0.723 (0.034)
Girls				
Y: OWO status @5 years (*N* = 981) (base model-0.64 [0.061])				
n = 1	0.675 (0.016)	0.722 (0.037)	0.731 (0.029)	0.743 (0.019)
n = 2	0.744 (0.025)	0.785 (0.028)	0.784 (0.022)	0.801 (0.008)
Y: OWO status @6 years (*N* = 924) (base model-0.617 [0.048])				
n = 1	0.663 (0.013)	0.71 (0.032)	0.711 (0.022)	0.732 (0.038)
n = 2	0.731 (0.021)	0.763 (0.019)	0.769 (0.034)	0.782 (0.015)
Y: OWO status @7 years (*N* = 864) (base model-0.635 [0.036])				
n = 1	0.637 (0.046)	0.691 (0.051)	0.69 (0.03)	0.719 (0.044)
n = 2	0.714 (0.041)	0.736 (0.051)	0.761 (0.041)	0.764 (0.041)
Y: OWO status @8 years (*N* = 765) (base model-0.634 [0.028])				
n = 1	0.637 (0.031)	0.691 (0.031)	0.697 (0.028)	0.72 (0.034)
n = 2	0.701 (0.018)	0.725 (0.035)	0.751 (0.033)	0.754 (0.034)
Y: OWO status @9 years (*N* = 650) (base model-0.612 [0.044])				
n = 1	0.627 (0.025)	0.689 (0.054)	0.67 (0.04)	0.704 (0.035)
n = 2	0.682 (0.031)	0.715 (0.047)	0.717 (0.041)	0.731 (0.031)
Y: OWO status @10 years (*N* = 572) (base model-0.648 [0.047])				
n = 1	0.619 (0.056)	0.669 (0.084)	0.69 (0.05)	0.714 (0.057)
n = 2	0.676 (0.044)	0.698 (0.089)	0.734 (0.047)	0.726 (0.061)
Boys				
Y: OWO status @5 years (*N* = 948) (base model-0.625 [0.046])				
n = 1	0.687 (0.03)	0.735 (0.038)	0.736 (0.029)	0.749 (0.039)
n = 2	0.733 (0.033)	0.778 (0.027)	0.771 (0.031)	0.793 (0.031)
Y: OWO status @6 years (*N* = 878) (base model-0.634 [0.046])				
n = 1	0.666 (0.04)	0.718 (0.053)	0.72 (0.026)	0.734 (0.052)
n = 2	0.712 (0.066)	0.758 (0.051)	0.75 (0.059)	0.771 (0.045)
Y: OWO status @7 years (*N* = 838) (base model-0.618 [0.048])				
n = 1	0.639 (0.03)	0.693 (0.035)	0.684 (0.021)	0.705 (0.036)
n = 2	0.692 (0.054)	0.737 (0.043)	0.72 (0.045)	0.744 (0.039)
Y: OWO status @8 years (*N* = 781) (base model-0.625 [0.033])				
n = 1	0.617 (0.027)	0.659 (0.054)	0.657 (0.02)	0.67 (0.054)
n = 2	0.677 (0.041)	0.711 (0.038)	0.708 (0.044)	0.718 (0.041)
Y: OWO status @9 years (*N* = 678) (base model-0.641 [0.057])				
n = 1	0.609 (0.04)	0.653 (0.038)	0.664 (0.037)	0.68 (0.046)
n = 2	0.659 (0.053)	0.704 (0.033)	0.705 (0.052)	0.723 (0.036)
Y: OWO status @10 years (*N* = 586) (base model-0.63 [0.066])				
n = 1	0.592 (0.034)	0.642 (0.04)	0.65 (0.055)	0.672 (0.053)
n = 2	0.638 (0.061)	0.693 (0.029)	0.687 (0.043)	0.703 (0.038)

Significance level is 5%. As this is 95% CI, and bold value means it does not contain/cross the value 1, so the P value is strictly less than 0.05.

*Crude model: static BMI category or BMIPCT trajectory as predictor only.

†Adjusted model: static BMI category and BMIPCT trajectory and covariates included as predictors.

‡Base model: conventional covariates are included as predictors, which include all characteristics listed in Table 1.

AUC, area under the ROC curve; BMI, body mass index; BMIPCT, body mass index percentile; CV, cross validation; OWO, overweight or obesity; SD, standard deviation.

better predictive performance compared to traditional clinical data or static BMI categories also suggests that early childhood BMIPCT trajectories can be used to help clinicians identify those children most at risk for OWO in middle childhood. Further studies are encouraged to investigate the predictive power of BMIPCT trajectory on other health outcomes, such as diabetes and high blood pressure. This will enable a more comprehensive understanding of the clinical utility of early childhood BMIPCT

trajectories in identifying children at risk of adverse health outcomes in later life.

To the best of our knowledge, this is the first study to identify BMIPCT trajectories at different age windows across the childhood period (from birth to age *n* years, *n* = 1, 2, …,18) in this underrepresented population. It is noted that Cao *et al*[15] also identified four BMIPCT trajectories from birth to age 18 years in a subset of this population by a combination of *K*-means clustering (*K* = 2) and principal component analysis. However, they imputed missing BMI measurements by averaging the last and next observed values and did not consider any time limit for the imputation range. We addressed this limitation by the proposed LOWESS-TLS, which increased our statistical power by incorporating 3029 samples (with missing records) and doubling the available data points, which is not typically observed in other similar studies.[13,15,21,25] It is noted that, to the authors' knowledge, this appears to be the first attempt to assess the predictive power of early childhood (age 1 or 2 years) BMIPCT trajectories on OWO in middle childhood (age 5–10 years).

There are several limitations in our study. We only employed *K*-means clustering and LCGA, even though other methods could also be used for trajectory analysis (eg, latent class analysis[52] and latent class growth mixture modeling[19]). Further research is needed to ascertain more generalized conclusions as our research focused on the BBC. Additionally, although we applied novel imputation methods and our imputed data were statistically consistent with the raw data, there are still other imputation methods, such as multiple imputation,[53] which can substantially reduce missing data bias, introduce appropriate random errors, and avoid significant loss of sample size.

In conclusion, we have demonstrated how a novel longitudinal data imputation method that utilizes clinically relevant time intervals and *K*-means clustering can be used to identify BMIPCT growth trajectories. We further investigated the associations of maternal and infant characteristics with BMIPCT trajectories and the predictive power of the trajectories on long-term OWO risk. These findings can help to advance early life risk assessment and prediction, enabling targeted interventions to prevent obesity, a major risk factor for noncommunicable diseases,[54] at early developmental windows.

Acknowledgments

We thank Tami Bartell, MPH, for her helpful review and editing of this manuscript.

Funding

The Boston Birth Cohort (the parent study) is supported in part by the National Institutes of Health (NIH) grants (R21ES011666, 2R01HD041702, R21HD066471, R01HD086013, R01HD098232, R01ES031272, R01ES031521, and U01ES034983); and the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) (UT7MC45949). Dr. Guoying Wang is also supported by grant R03ES029594 from the NIH/National Institute of Environmental Health Science. This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by any funding agencies.

Author Contributions

Conceptualization: WH, TLC, MV, BO, LL, TI, XW; data curation: WH, XH, GW, XW; methodology and formal analysis: WH, AYM, LL, TI, XW; funding acquisition: XW; project administration, resources; software: TI, XW; supervision: LL, TI, XW; writing - original draft: WH; writing - review & editing: all authors contributed to the critical review and revision of the manuscript and approved the final manuscript. WH, XH, LL, and XW have full access to the data and verified the underlying analyses and results.

Conflicts of Interest

The authors declare no conflict of interests.

Disclaimer

The opinions and assertions expressed herein are those of the author(s) and do not reflect the official policy or position of the Uniformed Services University of the Health Sciences or the Department of Defense.

References

[1] CDC. Overweight & obesity: defining childhood obesity. Centers for Disease Control and Prevention. Available from: https://www.cdc.gov/obesity/childhood/defining.html. Published 2018. [Accessed April 24, 2020].
[2] Sahoo K, Sahoo B, Choudhury AK, et al. Childhood obesity: causes and consequences. J Family Med Prim Care 2015;4(2):187–192. doi:10.4103/2249-4863.154628.
[3] Young TK, Dean HJ, Flett B, et al. Childhood obesity in a population at high risk for type 2 diabetes. J Pediatr 2000;136(3):365–369. doi:10.1067/mpd.2000.103504.
[4] Bao W, Srinivasan SR, Wattigney WA, et al. Persistence of multiple cardiovascular risk clustering related to syndrome X from childhood to young adulthood. The Bogalusa Heart Study. Arch Intern Med 1994;154(16):1842–1847. doi:10.1001/archinte.1994.00420160079011.
[5] Han JC, Lawlor DA, Kimm SY. Childhood obesity. Lancet 2010;375(9727):1737–1748. doi:10.1016/S0140-6736(10)60171-7.
[6] Deckelbaum RJ, Williams CL. Childhood obesity: the health issue. Obes Res 2001;9(Suppl 4):239S–243S. doi:10.1038/oby.2001.125.
[7] CDC. Childhood Obesity Facts. Available from: https://www.cdc.gov/obesity/data/childhood.html#Prevalence. Published 2021. [Accessed July 20, 2022].
[8] Franks PW, Hanson RL, Knowler WC, et al. Childhood obesity, other cardiovascular risk factors, and premature death. N Engl J Med 2010;362(6):485–493. doi:10.1056/NEJMoa0904130.
[9] Ochoa MC, Moreno-Aliaga MJ, Martinez-Gonzalez MA, et al. Predictor factors for childhood obesity in a Spanish case-control study. Nutrition 2007;23(5):379–384. doi:10.1016/j.nut.2007.02.004.
[10] Simmonds M, Llewellyn A, Owen CG, et al. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. Obes Rev 2016;17(2):95–107. doi:10.1111/obr.12334.
[11] Huang W, Igusa T, Wang G, et al. In-utero co-exposure to toxic metals and micronutrients on childhood risk of overweight or obesity: new insight on micronutrients counteracting toxic metals. Int J Obes (Lond) 2022;46(8):1435–1445. doi:10.1038/s41366-022-01127-x.
[12] Haga C, Kondo N, Suzuki K, et al. Developmental trajectories of body mass index among Japanese children and impact of maternal factors during pregnancy. PLoS One 2012;7(12):e51896e51896. doi:10.1371/journal.pone.0051896.
[13] Mattsson M, Maher GM, Boland F, et al. Group-based trajectory modelling for BMI trajectories in childhood: a systematic review. Obes Rev 2019;20(7):998–1015. doi:10.1111/obr.12842.
[14] Twisk J, Hoekstra T. Classifying developmental trajectories over time should be done with great caution: a comparison between methods. J Clin Epidemiol 2012;65(10):1078–1087. doi:10.1016/j.jclinepi.2012.04.010.
[15] Cao T, Zhao J, Hong X, et al. Cord blood metabolome and BMI trajectory from birth to adolescence: a prospective birth cohort study on early life biomarkers of persistent obesity. Metabolites 2021;11(11):739. doi:10.3390/metabo11110739.

[16] Smith AJ, O'Sullivan PB, Beales DJ, et al. Trajectories of childhood body mass index are associated with adolescent sagittal standing posture. Int J Pediatr Obes 2011;6(2-2):e97–106. doi:10.3109/17477166.2010.530664.
[17] Bhattacharya M, Ehrenthal D, Shatkay H. Identifying growth-patterns in children by applying cluster analysis to electronic medical records. 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Belfast, UK, 2014, pp. 348–351. doi:10.1109/BIBM.2014.6999183.
[18] Fuemmeler BF, Yang C, Costanzo P, et al. Parenting styles and body mass index trajectories from adolescence to adulthood. Health Psychol 2012;31(4):441–449. doi:10.1037/a0027927.
[19] Berlin KS, Parra GR, Williams NA. An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. J Pediatr Psychol 2014;39(2):188–203. doi:10.1093/jpepsy/jst085.
[20] Nanishi M, Fujiogi M, Stevenson M, et al. Association of growth trajectory profiles with asthma development in infants hospitalized with bronchiolitis. J Allergy Clin Immunol Pract 2022;10(3):723–731.e5. doi:10.1016/j.jaip.2021.11.001.
[21] Aris IM, Chen LW, Tint MT, et al. Body mass index trajectories in the first two years and subsequent childhood cardio-metabolic outcomes: a prospective multi-ethnic Asian cohort study. Sci Rep 2017;7(1):8424. doi:10.1038/s41598-017-09046-y.
[22] Boyer BP, Nelson JA, Holub SC. Childhood body mass index trajectories predicting cardiovascular risk in adolescence. J Adolesc Health 2015;56(6):599–605. doi:10.1016/j.jadohealth.2015.01.006.
[23] Stuart B, Panico L. Early-childhood BMI trajectories: evidence from a prospective, nationally representative British cohort study. Nutr Diabetes 2016;6:e198. doi:10.1038/nutd.2016.6.
[24] Giles LC, Whitrow MJ, Davies MJ, et al. Growth trajectories in early childhood, their relationship with antenatal and postnatal factors, and development of obesity by age 9 years: results from an Australian birth cohort study. Int J Obes (Lond) 2015;39(7):1049–1056. doi:10.1038/ijo.2015.42.
[25] Bichteler A, Gershoff ET. Identification of Children's BMI trajectories and prediction from weight gain in infancy. Obesity (Silver Spring) 2018;26(6):1050–1056. doi:10.1002/oby.22177.
[26] Demment MM, Haas JD, Olson CM. Changes in family income status and the development of overweight and obesity from 2 to 15 years: a longitudinal study. BMC Public Health 2014;14:417. doi:10.1186/1471-2458-14-417.
[27] Wang G, DiBari J, Bind E, et al. Association between maternal exposure to lead, maternal folate status, and intergenerational risk of childhood overweight and obesity. JAMA Netw Open 2019;2(10):e191 2343e1912343. doi:10.1001/jamanetworkopen.2019.12343.
[28] Pearson C, Bartell T, Wang G, et al. Boston Birth Cohort profile: rationale and study design. Precision Nutr 2022;1(2):e00011. doi:10.1016/S2214-109X(18)30451-0.
[29] Wang G, Divall S, Radovick S, et al. Preterm birth and random plasma insulin levels at birth and in early childhood. JAMA 2014;311(6):587–596. doi:10.1001/jama.2014.1.
[30] Mao G, Nachman RM, Sun Q, et al. Individual and joint effects of early-life ambient exposure and maternal prepregnancy obesity on childhood overweight or obesity. Environ Health Perspect 2017;125(6):067005. doi:10.1289/EHP261.
[31] CDC. WHO growth chart. Available from: https://www.cdc.gov/growthcharts/who_charts.htm. Published 2006. Updated September 9, 2010. [Accessed March 3, 2022].
[32] CDC (Centers for Disease Control and Prevention). CDC growth chart. Available from: https://www.cdc.gov/growthcharts/. Published 2000. Updated November 26, 2013. [Accessed April 14, 2020].
[33] Hong X, Wang G, Liu X, et al. Gene polymorphisms, breast-feeding, and development of food sensitization in early childhood. J Allergy Clin Immunol 2011;128(2):374–81.e2. doi:10.1016/j.jaci.2011.05.007.
[34] Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple imputation in clinical epidemiological research. Clin Epidemiol 2017;9:157–166. doi:10.2147/CLEP.S129785.
[35] Engels JM, Diehr P. Imputation of missing longitudinal data: a comparison of methods. J Clin Epidemiol 2003;56(10):968–976. doi:10.1016/s0895-4356(03)00170-7.
[36] Hartigan JA, Wong MA. Algorithm AS 136: a k-means clustering algorithm. J R Stat Soc Ser C Appl Stat 1979;28(1):100–108. doi:10.2307/2346830.
[37] Scott Shaobing C, Gopalakrishnan PS. Clustering via the Bayesian information criterion with applications in speech recognition. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181), Seattle, WA, 1998, pp. 645–648 vol.2, doi:10.1109/ICASSP.1998.675347.
[38] Zhao Q, Hautamaki V, Fränti P. Knee Point Detection in BIC for Detecting the Number of Clusters. Paper presented at: Advanced Concepts for Intelligent Vision Systems: 10th International Conference, ACIVS 2008, Juan-les-Pins, France, October 20-24, 2008. Proceedings 10, pp. 664–673. Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-88458-3_60.
[39] Muthén B. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan D, ed. The SAGE Handbook of Quantitative Methodology for the Social Sciences. SAGE Publications, Inc.; 2004.
[40] Halfon N, Forrest CB, Lerner RM, Faustman EM, Tullis E, Son J. Introduction to the handbook of life course health development. In: Halfon N, Forrest CB, Lerner RM, Faustman EM, eds. Handbook of Life Course Health Development. Cham (CH); 2018. p. 1–16.
[41] Shapiro ALB, Moore BF, Sutton B, et al. In utero exposure to maternal overweight or obesity is associated with altered offspring brain function in middle childhood. Obesity (Silver Spring) 2020;28(9):1718–1725. doi:10.1002/oby.22908.
[42] CDC. Middle Childhood (9-11 years of age). Available from: https://www.cdc.gov/ncbddd/childdevelopment/positiveparenting/middle2.html. Published 2021. Updated September 23, 2021. [Accessed December 22, 2021].
[43] Whitaker RC, Wright JA, Pepe MS, et al. Predicting obesity in young adulthood from childhood and parental obesity. N Engl J Med 1997;337(13):869–873. doi:10.1056/NEJM199709253371301.
[44] Singh AS, Mulder C, Twisk JWR, et al. Tracking of childhood overweight into adulthood: a systematic review of the literature. Obes Rev 2008;9(5):474–488. doi:10.1111/j.1467-789x.2008.00475.x.
[45] Wanyu Huang TI, Guoying W, Jessie P, et al. In-utero co-exposure to toxic metals and micronutrients on childhood risk of overweight or obesity: new insight on micronutrients counteracting toxic metals. Int J Obes 2022;46(8):1435–1445. doi:10.1038/s41366-022-01127-x.
[46] Pryor LE, Tremblay RE, Boivin M, et al. Developmental trajectories of body mass index in early childhood and their risk factors: an 8-year longitudinal study. Arch Pediatr Adolesc Med 2011;165(10):906–912. doi:10.1001/archpediatrics.2011.153.
[47] Wang G, Johnson S, Gong Y, et al. Weight gain in infancy and overweight or obesity in childhood across the gestational spectrum: a prospective birth Cohort study. Sci Rep 2016;6:29867. doi:10.1038/srep29867.
[48] Tremblay MS, Willms JD. Is the Canadian childhood obesity epidemic related to physical inactivity? Int J Obes Relat Metab Disord 2003;27(9):1100–1105. doi:10.1038/sj.ijo.0802376.
[49] O'Dea JA, Caputi P. Association between socioeconomic status, weight, age and gender, and the body image and weight control practices of 6- to 19-year-old children and adolescents. Health Educ Res 2001;16(5):521–532. doi:10.1093/her/16.5.521.
[50] Suzuki K, Ando D, Sato M, et al. The association between maternal smoking during pregnancy and childhood obesity persists to the age of 9-10 years. J Epidemiol 2009;19(3):136–142. doi:10.2188/jea.je20081012.
[51] Toschke AM, Koletzko B, Slikker W, Jr, et al. Childhood obesity is associated with maternal smoking in pregnancy. Eur J Pediatr 2002;161(8):445–448. doi:10.1007/s00431-002-0983-z.
[52] Formann AK, Kohlmann T. Latent class analysis in medical research. Stat Methods Med Res 1996;5(2):179–211. doi:10.1177/096228029600500205.
[53] Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393. doi:10.1136/bmj.b2393.
[54] WHO. Obesity and overweight. Available from: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. Published 2021. [Accessed July 21, 2022].

How to cite this article: Huang W, Meir AY, Olapeju B, Wang G, Hong X, Venkataramani M, Cheng TL, Igusa T, Liang L, Wang X. Defining longitudinal trajectory of body mass index percentile and predicting childhood obesity: methodologies and findings in the Boston Birth Cohort. Precis Nutr 2023;2(2):e00037. doi: 10.1097/PN9.0000000000000037