

A framework for conducting time-varying genome-wide association studies: An application to body mass index across childhood in six multiethnic cohorts

Kimberley Burrows^{1,2*}, Anni Heiskala^{3*}, Jonathan P. Bradfield^{4,5*}, Zhanna Balkhiyarova^{6,7,8*}, Lijiao Ning⁹, Mathilde Boissel⁹, Yee-Ming Chan^{10,11}, Philippe Froguel^{9,12}, Amelie Bonnefond^{9,12}, Hakon Hakonarson⁴, Alessandro Couto Alves¹³, Deborah A Lawlor^{1,2}, Marika Kaakinen^{6,7,14}, Marjo-Riitta Järvelin^{3,15,16}, Struan F.A. Grant^{4,17,18,19,20}, Kate Tilling¹, Inga Prokopenko^{6,7,†}, Sylvain Sebert^{3,†}, Mickaël Canouil^{9,†}, Nicole M Warrington^{21,22,†}, on behalf of the Early Growth Genetics (EGG) Consortium

* Contributed equally as first authors

† Contributed equally as senior authors

Corresponding authors:

Nicole M Warrington (n.warrington@uq.edu.au)

Sylvain Sebert (Sylvain.Sebert@oulu.fi)

Author affiliations:

¹ MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK.

² Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

³ Research Unit of Population Health, University of Oulu, Oulu, Finland

⁴ Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

⁵ Quantinuum Research LLC, Wayne, PA, USA

⁶ Department of Clinical and Experimental Medicine, School of Biosciences and Medicine, University of Surrey, Guildford, UK

⁷ People-Centred Artificial Intelligence Institute, University of Surrey, Guildford, UK

⁸ Section of Metabolism, Digestion and Reproduction, Department of Medicine, Imperial College London, London, UK

⁹ Univ Lille, INSERM/CNRS UMR1283/8199, EGID, Institut Pasteur de Lille, Lille University Hospital, Lille, France

¹⁰ Division of Endocrinology, Department of Pediatrics, Boston Children's Hospital

¹¹ Department of Pediatrics, Harvard Medical School

¹² Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom

¹³ School of Biosciences and Medicine, University of Surrey, Guildford, UK

¹⁴ Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK

¹⁵ MRC Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom

¹⁶ Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, United Kingdom

¹⁷ Center for Spatial and Functional Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

¹⁸ Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

¹⁹ Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

²⁰ Divisions of Human Genetics and Endocrinology and Diabetes, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

²¹ Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

²² Frazer Institute, University of Queensland, Brisbane, Australia

Abstract

Genetic effects on changes in human traits over time are understudied and may have important pathophysiological impact. We propose a framework that enables data quality control, implements mixed models to evaluate trajectories of change in traits, and estimates phenotypes to identify age-varying genetic effects in genome-wide association studies (GWASs). Using childhood body mass index (BMI) as an example, we included 71,336 participants from six cohorts and estimated the slope and area under the BMI curve within four time periods (infancy, early childhood, late childhood and adolescence) for each participant, in addition to the age and BMI at the adiposity peak and the adiposity rebound. GWAS on each of the estimated phenotypes identified 28 genome-wide significant variants at 13 loci across the 12 estimated phenotypes, one of which was novel (in *DAOA*) and had not been previously associated with childhood or adult BMI. Genetic studies of changes in human traits over time could uncover novel biological mechanisms influencing quantitative traits.

Introduction

Genome-wide association studies (GWAS) have been informative over the past two decades in extending our knowledge of the genetic architecture of common complex traits and diseases. The vast majority of GWAS have concentrated on cross-sectional phenotypes (i.e., one measure per person per study). However, many human traits change over time, and there may be a genetic component underlying this dynamic process of change in the trait (see for example, E. N. Smith et al. 2010; Warrington et al. 2015; A. D. Smith et al. 2016; Gouveia et al. 2019; 2021; Wendel et al. 2021; Venkatesh et al. 2023). Therefore, studying trait trajectories is increasingly important to uncover novel loci beyond those found from GWAS of cross-sectional traits.

Linear mixed models (LMMs) are often used to assess age- (or time) varying exposure-outcome associations (Molenberghs and Verbeke 2000). This statistical model summarises repeated measures data into an average trajectory across the sample (i.e. the fixed effects) as well as the individual variations around this average for each participant within the sample (i.e. the random effects). LMMs can be used to explore age- (time) varying effects of genetic variants on outcomes. Indeed, there are examples of using LMMs to estimate the association between selected SNPs and trajectories of change in phenotypes (see for example, Sovio et al. 2011; Mook-Kanamori et al. 2011; Howe et al. 2013; Brand et al. 2019; Elhakeem et al. 2023). However, scaling this approach up to conduct GWAS on longitudinal change in a trait can be extremely computationally intensive; for example, analysing 2.5 million SNPs and longitudinal BMI data in 7,916 individuals from the Avon Longitudinal Study of Parents and Children (ALSPAC) took approximately 1,440 hours to complete (Warrington et al. 2015), in comparison to a few hours using standard software with a single measure per individual.

There are current statistical methods available to utilise repeated measures of a phenotype within an individual and facilitate the detection of genetic variants that have age- (time) varying effects when the phenotype changes linearly over time. This includes methods that use a two-stage approach, in which (i) a LMM is fit to the repeated measures data and (ii) the best linear unbiased predictors of each individual's trajectory are extracted and used as the outcome in the GWAS (Sikorska et al. 2013; Meirelles et al. 2013; Yuan et al. 2021). These two-step approaches reduce the computational challenge of fitting the LMM, allowing the GWAS to be conducted using standard software. As an alternative, Sikorska and colleagues (Sikorska et al. 2018) developed a method whereby the variances of the random intercept and slope, along with the variance of the residual, are estimated in a LMM without a genetic variant. This relies on the assumption that the

variances do not change when the genetic variant is included in the model. These estimated variances are then used in a large system of linear equations, which is solved to estimate the genetic effects at each SNP. These methods estimate the association between a genetic variant and change in a phenotype over time. In contrast, Ko and colleagues (2022) developed a method that estimates the genetic effect on within-subject variability that can be applied to biobank-scale repeated-measures data (Ko et al. 2022). All these methods described for longitudinal GWAS analyses are applicable for phenotypes that follow linear trajectories over time (Sikorska et al. 2018; Yuan et al. 2021; Ko et al. 2022). However, this is not realistic for many traits, particularly when investigating change in a phenotype over a long period of time.

An example of a trait with a non-linear trajectory is body mass index (BMI) across childhood. The BMI trajectory across childhood starts with a rapid increase soon after birth until the “adiposity peak” (AP) at approximately 9 months of age, followed by a gradual decline until the “adiposity rebound” (AR) around 4–6 years of age, followed by an increase again until the end of puberty and beyond (Rolland-Cachera et al. 1984). There is some evidence for age-varying genetic effects on BMI (Hardy et al. 2010; Elks et al. 2010; Sovio et al. 2011; Warrington et al. 2015); however, due to the lack of statistical methods for analysing non-linear trajectories on a GWAS scale, relatively little is known about the genetic determinants of the rate of change in BMI across early life.

The aim of this study is to develop a framework to conduct GWAS to detect age- (time) varying genetic effects of non-linear trajectories using standard GWAS software. Due to restrictions in sharing individual participant data, this framework will be applied to individual participant data within each cohort and then GWAS summary statistics will be meta-analysed. The framework comprised of the following four procedures: (1) apply an algorithm to quality control the longitudinal data to ensure that only the most likely outliers are excluded based on within- and between-individual comparisons; (2) specify an appropriate model for a nonlinear growth trajectory using longitudinal data in a range of different cohorts; (3) ensure the chosen model is correctly parameterized; (4) estimate phenotypes that summarise the trajectory for subsequent GWAS analysis. The resulting GWAS summary statistics can be used in downstream analyses, such as genetic correlation and causal modelling. We describe an example standardised protocol via an easy-to-follow R package, called Early Growth Genetics Longitudinal Analysis (*EGGLA*), to perform each of these steps for BMI across childhood and provide a harmonised, reproducible set of GWAS summary statistics for further downstream analysis.

Results

Our analysis included participants from six population-based cohorts (Table 1, Supplementary Table 1 and Supplementary text): 1) the Avon Longitudinal Study of Parents and Children (ALSPAC) (Boyd et al. 2013; Fraser et al. 2013), 2) the European subset of Children's Hospital of Philadelphia (CHOP) (Connolly et al. 2020) 3) the African American subset of CHOP (Connolly et al. 2020), 4) the Northern Finland Birth Cohort 1966 (NFBC1966) (Nordström et al. 2021), 5) the Northern Finland Birth Cohort 1986 (NFBC1986) (Järvelin et al. 1997), and 6) OBésité de l'Enfant (OBE) (Meyre et al. 2004; 2009). We included measures of BMI at multiple times across early life, ranging from two weeks after birth to late adolescence (18 years for all cohorts except OBE, which had data until 16 years).

Using a published algorithm to clean longitudinal data ensures that only the most likely outliers and errors are excluded

We employed a unified approach to data cleaning across cohorts by using the *growthcleanr* (Daymont et al. 2017) R package, which flags duplicates and implausible values for exclusion. After data cleaning, we excluded between 4.2% and 16% of BMI measures within each cohort, most of which were excluded due to missing height or weight or duplicated measures (Supplementary Table 2). The final analysis for the growth modelling comprised 34,818 females (ranging from 308 to 10,814 per cohort) and 36,518 males (ranging from 252 to 12,002 per cohort).

Developing an R package to conduct and compare nonlinear growth models across cohorts from different generations and of different ethnicities

We developed an analysis framework to fit nonlinear growth models for GWAS and associated R package (<https://m.canouil.dev/eggla/articles/eggla.html>; Canouil et al. 2024; R Core Team 2022) named the Early Growth Genetics (EGG) Longitudinal Analysis (*EGGLA*) framework. The *EGGLA* R package provides four unified protocols to facilitate the analysis framework, including model diagnostics, model selection, running the chosen LMM, and estimating specific phenotypes. The R package was developed to standardise the analysis framework, allowing all six participating cohorts to provide a harmonised, reproducible set of GWAS summary statistics for further downstream analysis. The *EGGLA* protocols specific to longitudinal modelling of BMI are outlined in Supplementary Figure 1 and the Supplementary Text. The *EGGLA* model diagnostics protocol cycles through each of the three selected LMMs (linear spline, cubic spline, and cubic slope functions for age) with various complexities of random effects and correlation structure (no

structure and continuous autoregressive correlation structure of order 1 (CAR(1)); see methods for full description of the LMMs). Several reports are output at this stage to inform selection of the most appropriate model to characterise change in the phenotype (BMI in our illustrative application; see Supplementary Table 3 and Supplementary Text). To select our preferred model, we compared the model fit based on any convergence issues, performance metrics and visual inspection of the predicted curves.

Several models either failed to converge or presented warning messages within each cohort (more issues were seen in cohorts with higher sample size and a larger number of repeated measures per individual; Table 2 and Supplementary Table 3). In each cohort, at least eight of the sixteen models converged without any issues (Supplementary Table 3). Specifying a CAR(1) correlation structure seemed to cause the most problems with model convergence across the cohorts/sexes, particularly for the models with a cubic slope or linear spline function for age in the fixed effects.

We used the intraclass correlation coefficient (ICC), conditional R^2 , root mean squared error (RMSE) and residual standard deviation (SD) performance metrics to define the best fitting model. In ALSPAC, females in NFBC1966, NFBC1986, and OBE the best model included a cubic spline function for age in both the fixed and random effects and no correlation structure (Table 2 and Supplementary Table 3). For the males in NFBC1966, the best fitting model included a cubic spline function for age in the fixed effects, a cubic slope in the random effects and no correlation structure specified. The African American subset of CHOP was also best fit by the cubic spline function for age in the fixed effects and no correlation structure; however, the best-fitting random effects (cubic slope, cubic spline or linear spline function for age) was different for different performance metrics. Finally, the best fitting model in the European American subset of CHOP had a cubic slope in the fixed and random effects, with the CAR(1) correlation structure producing the best fit for three of the performance metrics (ICC, conditional R^2 , and SD) but no correlation structure producing the best fit for RMSE. Although the cubic spline in both fixed and random effects produced the most favourable performance metrics for most cohorts (Table 2 and Supplementary Table 3), the model fit with warning messages in the CHOP European Americans and males in NFBC1966 suggesting that the next best performing models (that converge without issues for all cohorts and sexes) should be considered. Therefore, our preferred model included a cubic spline function for age in the fixed effects, cubic slope in the random effects and no specified correlation structure. This model is described using the following equation:

$$\log_e(BMI_{it}) = \beta_0 + \sum_{j=1}^3 \beta_j Age_{it}^j + \sum_{k=1}^3 \beta_{k+3} (Age_{it} - \kappa_k)_+^3 + b_{0i} + \sum_{j=1}^3 b_{ji} Age_{it}^j + \varepsilon_{it}$$

Where the fixed parameters are represented by $\beta_0, \beta_1, \beta_2, \dots, \kappa_k$ is the k^{th} knot and:

$$(t - \kappa_k)_+ = \begin{cases} 0 & \text{if } t \leq \kappa_k \\ t - \kappa_k & \text{if } t > \kappa_k \end{cases}$$

The parameter estimates for the random effects are represented by $b_{0i}, b_{1i}, b_{2i}, \dots$, and ε_{it} are the error terms (see methods for further description on the model specification).

Parameterising the preferred model by refining the knot placement for the cubic spline function, ensuring a better fit to the features of the BMI trajectory

Because our preferred model included a cubic spline function for age, we sought to further optimize the knot placement (i.e. ages at which there is a change in slope) to ensure our model reflected the underlying BMI trajectory appropriately. We started with knot points at two, eight and 12 years based on previous research (Warrington et al. 2015), which modelled BMI data from one to 17 years of age. However, our modelling of slightly younger ages could require different knot points. To this end, we ran models incrementing the first knot point by half year increments and second knot point by yearly increments; we decided to not move the third knot point as there is less change in the BMI trajectory after the adiposity rebound (i.e. by 12 years of age) and both our study and Warrington *et al.* included data across this time period (i.e. 12-17 years). In addition to the previous metrics used for selecting a preferred model, we estimated the age and BMI at the AP and AR to help us distinguish between models (see methods for description on how the AP and AR were estimated).

Previous studies exploring the age of AP in European populations report an average age at approximately nine months (Cole, Freeman, and Preece 1995). In all the cohorts in this study, later placement of the first knot (at age one, one and a half, or two years) resulted in an increase in the estimated average age at AP (Supplementary Figure 2). The average age at AP when applying the first knot at one year of age across all included cohorts was 0.75 years (SD 0.05; Table 1 and Supplementary Table 1), whereas it was over one year of age when the first knot was at two years of age. In addition, the performance metrics were improved across the cohorts when applying the first knot at one year of age compared to at two years of age (Supplementary Figure 2). We therefore chose to apply the first knot at one year of age for our final preferred model.

When comparing the model fit while moving the second knot (i.e. testing a knot at age six, seven, or eight years) we found there was very little variation in the performance metrics or in the estimated average age at AP and AR across the cohorts (Supplementary Figure 2). Therefore, we chose to keep knot two at our *a priori* age of eight years.

Figure 1 shows the average BMI trajectories predicted from the fixed effects of our chosen model (with knot points at one, eight and 12 years) for each of the six cohorts. The OBE cohort had a substantially higher average BMI and steeper trajectory throughout early life reflecting the obese children recruited into the cohort (the BMI trajectory in OBE is from 2 weeks to 16 years of age as there were very few observations beyond age 16 years). The African American subset of the CHOP cohort had an earlier AR and steeper BMI trajectory throughout childhood than the other cohorts. The model fit the BMI data well in all cohorts tested (Figure 2 and Supplementary Figure 3).

Estimate phenotypes that summarise the BMI trajectory for subsequent GWAS analysis

To estimate phenotypes from the BMI trajectory for GWAS analysis, we defined intervals of approximately linear change in BMI. We defined four time windows: infancy (two weeks - six months), early childhood (1.5 - 3.5 years), late childhood (6.5 - 10 years) and adolescence (12 - 17 years). We calculated a BMI trajectory for each individual within each cohort by combining the estimated fixed effects, which are shared by all subjects within a cohort, with the predicted random effects, which are specific to each individual. We subsequently estimated phenotypes within each time window from the individual-specific BMI trajectories, including slopes and area under the BMI curve (AUC), in addition to age and BMI at the AP and AR. We did not estimate the adolescent phenotypes (slope and AUC) in OBE as they only had data to age 16 years.

A summary of each of the estimated phenotypes is presented in Table 1 and Supplementary Table 1. The derived phenotypes illustrated the expected differences between the cohorts, indicating that our methods are generalizable to a range of cohorts. For example, individuals in the NFBC1966 and OBE cohorts had a steeper infancy slope (0.56 log BMI units per year for males in NFBC1966 and 0.54 log BMI units per year in OBE), in contrast to individuals in NFBC1986 (0.36 log BMI units per year in males). This is equivalent to a 75% change in BMI over the first year for males in NFBC1966, which equates to approximately 10kg/m² over the first year if BMI is 14kg/m² at two weeks of age (approximate BMI at two weeks of age across the cohorts

based on the height and weight data in Supplementary Table 1). Whereas in NFBC1986 males, it is equivalent to a 43% change in BMI, equating to approximately 6 kg/m² over the first year if BMI is 14 kg/m² at two weeks of age. This was also reflected in NFBC1966 and OBE having a higher BMI at the AP (for example, 18.22 kg/m² in males from NFBC1966 and 18.39 kg/m² from OBE) while NFBC1986 had a lower BMI at AP (17.75 kg/m² in males). As seen in Figure 1 and Table 1, the age at AR was earlier for OBE (for example, 2.40 years for males) than the other cohorts (average age of 3.83 - 5.67 years). The mean rate of growth was similar across the cohorts during late childhood and adolescence, with the mean growth rate ranging between 0.02 - 0.06 log BMI units per year (Table 1 and Supplementary Table 1), which is equivalent to a 2-6% change in BMI per year.

We investigated the correlation between each of the estimated phenotypes (Supplementary Table 4) and found high correlation ($r > 0.70$) between several phenotypes consistently across the cohorts. For example, the infant and early childhood slopes showed a positive correlation ($r \geq 0.78$) in both males and females in all cohorts. Both the infant ($r \geq 0.86$) and early childhood ($r \geq 0.73$) slopes were positively correlated with the age at AP. The late childhood slope was negatively correlated with the age at AR ($r \leq -0.84$), which indicates that individuals with an earlier AR have a steeper slope from 6.5 to ten years. These correlations differed in the OBE cohort that included only children with obesity; the correlation between late childhood slope and age at AR were small for males ($r = -0.09$) and females ($r = -0.10$). The AUC's generally showed strong positive correlations with the BMI at the AP and AR (infant AUC with BMI at AP, child AUC with BMI at both AP and AR and late child and adolescent AUC with BMI at AR). None of the estimated phenotypes were strongly correlated with the adolescent slope in any cohort (all $r < 0.7$).

Finally, all of the estimated phenotypes were associated with BMI at age 18 years (16 years in OBE, as BMI at 18 years was not available), and the strength of the association increased over age (Supplementary Figure 4). For example, the adolescent AUC explained the majority of the variance in BMI at 18 years (between 74-91% across the cohorts), in contrast to a small amount of variance explained by the infancy AUC (between 0-10%)

Meta-analysing the GWAS summary statistics and estimating genetic correlations

We conducted GWAS for each of the estimated phenotypes within each cohort, then combined the results from the GWAS in cohorts with individuals of European ancestry using fixed effects inverse-variance meta-analysis (OBE was excluded from the meta-analysis of the adolescent

phenotypes as they did not have data beyond age 16 years). Using the summary statistics from the meta-analysis, we estimated the SNP-based heritability of each of the estimated phenotypes and genetic correlation between phenotypes using linkage disequilibrium score regression (LDSC). The SNP-based heritability of the estimated phenotypes ranged from 6-28% (Figure 3), which is comparable to the SNP-based heritability of BMI across childhood (15-45% (Helgeland et al. 2022)) and adulthood (22%, (Yengo et al. 2018)). Estimates of SNP-based heritability for the infancy slope, adolescent slope and age at AP were less than 10%, with large standard errors, indicating that the genetic correlation estimates with these phenotypes may not be reliable; however, we have presented them for completeness. AUC in infancy was correlated with AUC in early childhood ($r_g=0.91$) and the BMI at both AP ($r_g=1$) and AR ($r_g=0.66$), but showed relatively low genetic correlations with all the other estimated phenotypes ($r_g < |0.43|$), indicating a unique genetic profile for BMI during infancy. The genetic correlation between AUCs in subsequent time periods were high; for example, the genetic correlation between AUC in infancy and early childhood was $r_g=0.91$, AUC in early childhood and late childhood $r_g=0.79$, and AUC in late childhood and adolescence $r_g=0.96$. This indicates that there could be a common set of genes related to BMI across time.

Using LDSC and publicly available summary statistics, we estimated the genetic correlation between our estimated phenotypes based on longitudinal data and childhood BMI at different ages (Helgeland et al. 2022) and adult BMI using cross-sectional data (Yengo et al. 2018). We found high genetic correlation between our estimated AUCs and childhood BMI measured cross-sectionally at the beginning and end of each age window (Supplementary Table 5). For example, the genetic correlation between early childhood AUC and BMI at age 1.5 years (the beginning of our early childhood time window) was 0.80 (SE=0.08), which was similar to the genetic correlation of 0.82 (SE=0.09) with BMI at age 3 years (near the end of our early childhood time window). In contrast, we estimated a moderate genetic correlation between our estimated slopes and childhood BMI (Supplementary Table 5), where the genetic correlation between early childhood slope and BMI at age 1.5 and 3 years of age was 0.13 (SE=0.09) and 0.39 (SE=0.10) respectively.

We identified 28 genome-wide significant ($P < 5 \times 10^{-8}$) variants at 13 loci associated with at least one of the 12 estimated phenotypes (Figure 4-6, Supplementary Table 6). The number of estimated phenotypes the loci were associated with ranged from one phenotype (*LEPR* associated with only Age AR and *DAOA* associated with only early childhood AUC) to seven phenotypes (*SEC16B* was associated with age and BMI at AR, age at AP, late childhood and

adolescent AUC and early and late childhood slope). Of these 13 loci, 12 have previously been identified in GWAS of adulthood BMI or obesity-related traits and 9 have been associated with childhood BMI-related traits. We identified one novel locus in the *DAOA* region on chromosome 13 that was most strongly associated with AUC in early childhood (1.5-3.5 years). Here, the A allele at rs79577162 (*DAOA*) reduces the AUC within this time-period (effect size=-0.02 log BMI years, $P=4\times 10^{-8}$). Across the other derived phenotypes, the A allele at rs79577162 also decreases BMI at the AP (effect size=-0.13kg/m², $P=2\times 10^{-6}$) and AR (effect size=-0.14kg/m², $P=5\times 10^{-6}$), and decreases the AUC across all time periods (effect size for infancy AUC=-0.003 log BMI years, $P=3\times 10^{-5}$; effect size for late childhood AUC=-0.04 log BMI years, $P=1\times 10^{-4}$; effect size for adolescent AUC=-0.06 log BMI years, $P=3\times 10^{-3}$); there was no evidence that the SNP impacts the slope across childhood (all $P>0.24$) or the age at the AP ($P=0.25$) or AR ($P=0.25$). Results were similar when OBE was excluded from the meta-analysis (Supplementary Figure 5). Twelve of the 28 the genome-wide significant variants identified in the European meta-analysis showed the same direction of effect in the CHOP African American subset and reached nominal significance ($P<0.05$) for at least one estimated phenotype (Supplementary Table 7), and a further 12 loci showed the same direction of effect ($P>0.05$).

The adolescent slope was not strongly associated with any region of the genome (Figure 5). The most significant locus was near *FAM120AOS* (rs11790060, $P=7\times 10^{-8}$) on chromosome 9, a region previously associated with change in BMI over early life (Warrington et al. 2015) (Supplementary Figure 6).

Variants near *SEC16B*, which have previously been shown to associate with adult BMI (Thorleifsson et al. 2009), were associated with the majority of our estimated phenotypes. For example, the T allele at rs509325 is associated with decreased age at the AP (effect size=-0.003 years, $P=1\times 10^{-8}$), and it is also associated with increased age at the AR (effect size=0.12 years, $P=5\times 10^{-16}$) and decreased BMI at the AR (effect size=-0.08 kg/m², $P=5\times 10^{-8}$). The T allele is also associated with decreased early childhood slope between 1.5 and 3.5 years (-0.0011 log BMI units per year, $P=2\times 10^{-11}$) and late childhood slope between 6.5 and 10 years (-0.0013 log BMI units per year, $P=1\times 10^{-17}$), resulting in a lower AUC in the subsequent time periods (i.e. -0.0315 log BMI years, $P=3\times 10^{-13}$ for late childhood AUC and -0.0670 log BMI years, $P=4\times 10^{-17}$ for adolescent AUC). Variants near *FTO* and *ADCY3* are similarly associated with a number of the estimated phenotypes.

We also investigated the association between 112 unique SNPs previously identified childhood BMI or obesity associated SNPs and our estimated phenotypes (Supplementary Table 8). The majority of the 112 SNPs showed directionally concordant effects between the published BMI traits and our estimated phenotypes (76/112 SNPs increased infancy slope, 86/112 increased early childhood slope, 84/112 increased late childhood slope, 83/112 increased infancy AUC, 97/112 increased early childhood AUC, 96/112 increased late childhood AUC, 96/112 increased adolescent AUC, 80/112 increased age at AP, 89/112 increased BMI at AP, 100/112 increased BMI at AR), with the exception of the adolescent slope (24/112). Given that an earlier age at AR is associated with higher BMI in later life, 82/112 SNPs associated with higher BMI traits were associated with earlier age at AR. At least half of the SNPs reached nominal significance ($P < 0.05$) for the childhood to adolescent traits, except for adolescent slope, for which only 18/112 SNPs reached nominal significance.

Discussion

We have developed a framework to perform non-linear growth modelling and conduct GWAS to detect age- (time) varying genetic effects of non-linear trajectories. To perform the analysis for BMI across childhood in multiple cohorts using our framework, we created an easy-to-use R package, called Early Growth Genetics Longitudinal Analysis (*EGGLA*). We consider childhood BMI an important and valuable example for testing our framework due to the established complex changes in adiposity that occur during the childhood, such as the AR and pubertal effects on BMI. We have shown that a LMM with a cubic spline function in the fixed effects and a cubic slope in the random effects, which is similar to a model used previously (Warrington et al. 2015), fit well in all the cohorts tested. We subsequently estimated phenotypes from this LMM and used them to conduct GWAS analyses. We identified 28 SNPs from 13 loci that associate with one or more of the estimated childhood BMI phenotypes, one of which was novel and has not been previously associated with childhood or adult BMI. These include 6 loci that were associated with the estimated AUCs, 6 loci that were associated with both change in BMI over time (estimated slopes) and the estimated AUCs and one that was associated with the age at AR but none of the AUCs or slopes. Although the majority of the identified loci had previously been identified in GWAS of BMI measured cross-sectionally in childhood and/or adulthood, our framework allows exploration of how the genetic effect changes over age (time), which is difficult to elucidate from cross-sectional GWAS analyses. With a larger sample size, this framework is likely to be useful to explore GWAS of time varying phenotypes to identify novel genetic associations that are relatively stable across large age periods (i.e. through the association with estimated AUCs) and those that vary with age (through the association with estimated slopes).

Given we have demonstrated our framework works well with the complex changes in BMI across childhood, we believe it is generalizable to longitudinal analysis of other traits. For example, the same procedure could be applied to height trajectories across childhood or changes in lean and fat mass. It could also be applied to other non-anthropometric traits, such as blood pressure or cholesterol trajectories across adulthood. The challenging aspect that will require further development for each trait of interest is defining an appropriate model to fit the repeated measures data that accurately describes the trajectory. However, once an appropriate model has been identified for the particular trait, then phenotypes can be estimated for subsequent GWAS analyses. We have provided the R package appropriate for modelling BMI trajectories across early life and details of the framework here so that it can be explored for use with other traits.

Our model fit the BMI data across a range of cohorts, including different ethnicities and different time points for data collection. For example, the final model had a similar fit in the two subsets of the CHOP cohort (the European American and African American), and the average BMI trajectories predicted from the model followed the expected growth patterns from previous research (Min et al. 2018), with the African American cohort having a higher BMI and faster rate of growth than the European American cohort. In the OBE cohort, there was very sparse data between the ages of 16 and 18 years; when we removed data after 16 years from the modelling, the model fit in OBE was comparable to the other cohorts. It is unknown how these models will perform in cohorts with data across a shorter age range; it may be necessary to incorporate alternative methods when cohorts are included with data across a different age ranges.

We propose using the AUC as one of the summary measures of the trait trajectories. The AUC for childhood BMI has been described as "the child's cumulative 'exposure' to excessive body weight" (Wen et al. 2012), and has been used in epidemiological studies, but we have found no evidence that it has been used previously for GWAS of any trait. Using AUC has potential benefits over analysing BMI in a cross-sectional manner as it is a combination of both baseline BMI and the incremental change in BMI over the time-period, whereas cross-sectional BMI would be the BMI at a given time point. Additionally, using the AUC rather than BMI at a single time point potentially increases statistical power to detect a genetic effect as the multiple BMI measurements used to estimate the AUC would average out any errors in the measurements and therefore reduce the variance attributable to measurement error. The AUC was more highly phenotypically and genetically correlated with BMI at the AP and AR, as well as at age 18, whereas the slopes were more highly correlated with the timing of the AP and AR. Consistent with the correlations, the SNP effects for the AUC were more directionally concordant with previously identified BMI and obesity loci than the SNP effects for the slope. This indicates, at least for BMI in early life, that genetic studies of the estimated slope parameters could uncover novel biology driving BMI in childhood that the current studies of childhood BMI have failed to identify. In contrast, genetic studies of the estimated AUC parameters would be likely to provide similar findings to the current genetic studies of childhood BMI; however, because the AUC parameters are estimated from the trajectory, measurement of BMI at the exact same time points across cohorts is not necessary, so using AUC allows incorporation of more cohorts to enhance sample sizes and statistical power for genetic studies.

Our analyses of BMI showed consistent results with previous literature, further validating our approach. First, the phenotypic correlations between the age and BMI at the AP and AR are similar to the correlations from a paediatric cohort with data collected between 1980-2008 (Wen et al. 2012). For example, Wen et al (Wen et al. 2012) estimated the correlation between BMI at the AP and AR to be 0.76, and ours ranged between 0.69-0.91 across the cohorts. Similarly, their correlation between age and BMI at the AR was -0.48, and ours ranged between -0.39 to -0.67. This indicates that the dependencies between the estimated phenotypes are consistent regardless of the underlying LMM fit to the repeated measures data. Second, the genetic correlations between the infancy AUC/early childhood AUC/BMI at AP and the other estimated phenotypes were relatively low, indicating that genetic loci associated with BMI in the first 3-4 years of life are likely to be different from those associated with BMI in later life. This observation is consistent with both a twin study that has shown that the genetic correlations differ between early and middle childhood (Silventoinen et al. 2022) and the MoBa study that show that the genetic correlation between childhood and adult BMI dramatically increases after the age of 5 years (Helgeland et al. 2022). Third, the SNP-based heritability of the estimated phenotypes ranged from 6-28% and is similar to the SNP-based heritability of cross-sectional BMI in childhood (15-45%, (Helgeland et al. 2022)) and adulthood (22%, (Yengo et al. 2018)). The low SNP-based heritability observed for the infancy slope (0.07 (SD=0.03)) and age at AP (0.09 (SD=0.03)) was also seen by Couto Alves and colleagues (Couto Alves et al. 2019) (-0.03 (SD=0.08) for age AP). This low SNP-based heritability could be due to a higher environmental component operating at this age, a genetic component that is not tagged by the SNPs on the GWAS array or a maternal genetic component that is independent of the child's genetic component. Alternatively, given the SNP-based heritability from cross-sectional BMI GWAS between 6 weeks of age and 1 year in the MoBa study ranged from 0.2-0.45 (Helgeland et al. 2022), it could indicate that our estimated phenotypes are not proxying the underlying BMI trajectory well through this age range. Therefore, further investigation into this low SNP-based heritability is warranted.

The effect estimates at identified SNPs in our GWAS are consistent with previously reported effects on childhood growth. For example, Warrington *et al.* identified rs1558902 at the *FTO* locus to be associated with change in BMI over childhood, which is in high LD with rs55872725 ($D'=1$, $r^2=1$) (Warrington et al. 2015). We found that the T allele at rs55872725 is associated with increased rate of BMI change from infancy to late childhood ($P<7\times 10^{-7}$), but not associated with adolescent slope ($P=0.129$), which is consistent with the pattern of association identified by Warrington *et al.* (Warrington et al. 2015). Additionally, the T allele is associated with lower AUC

in infancy ($P=0.002$), not associated with AUC in early childhood ($P=0.502$) and then associated with higher AUC in late childhood and adolescence ($P<4\times 10^{-14}$), which again is consistent with Warrington et al. where the genetic effect is associated with decreased BMI from 1-2 years of age, not associated with BMI from 3-5 years of age, then associated with increased BMI from 6 years of age onwards (Warrington et al. 2015). Therefore, although we have used different methods to identify and describe the genetic effects on growth, we are able to recapitulate what has previously been described, validating our proposed framework.

We have used our model to estimate the BMI and age at both the AP and AR. However, we acknowledge that the biological significance of these two markers remains unclear. It was initially thought that the relationship between age at AR and later risk of obesity was due to both the number and size of adipocytes increasing (Rolland-Cachera et al. 1984). Later studies propose that it could be related to an increase in the rate of lean mass rather than fat mass development (Campbell et al. 2011; Plachta-Danielzik et al. 2013). Cole (Cole 2004) suggests that it is a statistical phenomenon driven by both high centile and upward centile crossing, which are separately associated with an early rebound. Our findings here, and further application of our framework to change in fat and lean mass across childhood, could provide insights related to these different hypotheses.

There are several limitations to our approach. Firstly, the trajectories for each individual (i.e. the best linear unbiased predictors) are biased towards the average trajectory of the cohort (a property of LMMs known as 'shrinkage'), particularly when the individual has fewer repeated measures or when their trajectory is vastly different from others in the cohort (Robinson 1991). This could result in a biased estimate of the SNP effect on the trajectory in GWAS. Future work could incorporate approaches to adjust for the bias introduced by shrinkage (for example, methods used by Sikorska et al. 2013; Meirelles et al. 2013; Yuan et al. 2021) into our framework. Secondly, we have not accounted for the high phenotypic and genetic correlation between the estimated phenotypes (see Figure 3 and Supplementary Table 4) in the GWAS meta-analysis. Performing a multivariate meta-analysis accounting for the high correlation between the estimated phenotypes may reduce the overlap between the genetic signals seen. It may also increase the power to detect a genetic locus by leveraging information from the other correlated phenotypes, as seen in analyses using the MTAG software (Turley et al. 2018). Further research on the most appropriate multivariate meta-analysis method is required. Third, one of the advantages of using repeated measurements per individual is to increase the statistical power to detect genetic

associations when the SNP is included in the fixed effects part of the LMM due to the smaller residual error variance. However, given we are reducing the dimension of our dataset back to a single measure per person (for each GWAS analysis), it is unclear what impact this has on the statistical power. Finally, BMI as a measure of adiposity during infancy is not commonly used clinically, with ponderal index (weight divided by height³) being preferred. However, to model the trajectory from 2 weeks to 18 years of age we needed to use one measure consistently. Further research into the effect of different powers of height, as investigated by Stergiakouli and colleagues (Stergiakouli et al. 2014), would be of interest.

In conclusion, we have described a novel framework for conducting GWAS meta-analyses on longitudinal (repeated measures) phenotypes that have a non-linear trajectory over time. We provide an R package, *EGGLA*, to conduct these analyses for childhood BMI consistently across different cohorts. We have shown that the estimated phenotypes summarise the BMI growth trajectory across a range of cohorts and are able to detect genetic associations with known BMI-associated regions of the genome and to detect new genetic associations. Performing similar analyses across a wider range of cohorts with BMI measures across childhood may identify additional novel loci for change in BMI across early life. Identification of such loci will enable downstream analyses, such as genetic correlation and causal modelling, to investigate relationships between, for example, early life growth and developmental milestones, cognition and later life cardio-metabolic disease.

Methods

Overview

The Early Growth Genetics (EGG) Longitudinal Analysis (*EGGLA*) framework comprises four main components: application of a LMM to longitudinal data (in this case, BMI measurements between 2 weeks and 18 years of age), model diagnostics, model refinement, and finally GWAS, all provided through a series of functions defined within the *EGGLA* R package (Canouil et al. 2024; R Core Team 2022). The *EGGLA* R package was developed to provide a unified approach to harmonise analyses within six distinct cohorts. Although the package and accompanying documentation is specific to this longitudinal analysis with childhood BMI, it (and the described methods) serves as an exemplar for other consortium efforts to model other non-linear traits. The *EGGLA* R package is therefore publicly available on GitHub; it is also available as part of a Docker image, which provides all necessary tools for conducting LMM and GWAS analyses allowing the analyses to be run non-interactively through either Shell Command Language, Bourne-Again SHell, or interactively through R v4.2.0 (or greater). Further details of the application of the *EGGLA* R package to cohorts within the EGG consortium are described below.

Cohorts and study participants

We chose six cohorts to test our proposed methods, which include prospective general population-based birth cohorts from different generations and geographical locations (ALSPAC (Boyd et al. 2013; Fraser et al. 2013), NFBC1966 (Nordström et al. 2021) and NFBC1986 (Järvelin et al. 1997)), a retrospective cohort of children with obesity (OBE (Meyre et al. 2004; 2009)), and different ethnicities (individuals of African American and European American descent in CHOP (Connolly et al. 2020)).

For the growth modelling, we included all individuals with anthropometric data between the ages of two weeks and 18 years, except for OBE, which had very sparse data between 16 and 18 years, so after careful testing we excluded any measures after 16 years for the OBE cohort. Measures of height and weight from birth to two weeks were excluded to mitigate the effects of the weight drop arising after birth (DiTomasso and Cloud 2019). We excluded individuals who were part of a multiple birth (*i.e.*, twins, triplets). All individuals with available anthropometric data were included in the growth modelling, regardless of whether they had genetic data, to ensure a precise estimate of the average BMI trajectory within each cohort. Cohort-specific covariates were included in the growth modelling; for example, ALSPAC included a binary indicator of

measurement source (questionnaire vs. clinic or health visitor measurement) to allow for differential measurement error.

Data quality control

We applied an automated algorithm developed by Daymont and colleagues (Daymont et al. 2017), which is available as the R package, *growthcleanr* (<https://cran.r-project.org/web/packages/growthcleanr/index.html>). This package compares each measurement with a weighted moving average of the individual's other measurements to identify biologically implausible values in height and weight. The cleaning function flags potential data errors including: unit-switch errors (e.g., pounds recorded as kgs or height recorded as weight), very extreme values (i.e., z-score > 25), carried forwards (i.e., values identical over time for the same individual), duplicates (i.e., values recorded on the same day), large height absolute differences (i.e., a decrease in height by more than 3 cm in sequential measurements), single measurements and pairs (i.e., individual SD score for height was compared to their SD score for weight and vice-versa), error load (i.e., exclude individuals who have a substantial proportion of values flagged to be excluded), and finally moderate outliers (i.e., found by calculating the exponentially weighted moving averages for moderate and extreme outliers). We excluded any height or weight values that were flagged as potential data errors (between 4.2% and 16% of measures within each cohort; Supplementary Table 2). We have incorporated the *growthcleanr* quality control into our *EGGLA* R package.

After applying the cleaning protocol, measurements of height (m) and weight (kg) were used to derive BMI as $\text{weight}(\text{kg}) / \text{height}(\text{m})^2$. As BMI is skewed and heteroskedastic, a natural log transformation was applied before analysis.

Linear mixed modelling

LMM is one commonly used approach for overcoming some of the challenges in modelling longitudinal data. By selecting an appropriate function for age, the average trajectories of an outcome (i.e., average relationship between age and BMI) can be estimated as fixed effects in the LMM, while variation around this average on the individual level can be estimated as random effects (Laird and Ware 1982; Goldstein and Stavola 2010). Further details on LMM are given in the Supplementary Text.

Application of LMMs to BMI

We applied LMMs to model the trajectory of BMI, on a natural-log scale, over time from two weeks to 18 years of age (except OBE, which had data until 16 years). We fit three different functions for age to capture the non-linear slope of the $\log_e(\text{BMI})$ trajectory: 1) a cubic slope for age, 2) linear smoothing splines with knot points at key inflection points on the curve, and 3) cubic smoothing splines with knot points between inflection points in the curve. While we acknowledge that these three functions for age are not the only functions that could have been used to model the non-linearity in the trajectory, we chose these as they range from a simplistic function to model the non-linear trajectory (linear smoothing splines) to more complex function (cubic smoothing splines), but other functions may be more appropriate for comparison when investigating other traits. The linear smoothing splines were included as they have the benefit of being able to be used as the estimated slope phenotypes, rather than requiring estimated slopes to be derived separately. The three models are described by the equations below for individual i at time point t .

1. *LMM with cubic slope for age:*

$$\log_e(\text{BMI}_{it}) = \beta_0 + \sum_{j=1}^3 \beta_j \text{Age}_{it}^j + b_{0i} + \sum_{j=1}^3 b_{ji} \text{Age}_{it}^j + \varepsilon_{it} \quad (1)$$

2. *LMM with linear smoothing splines:*

$$\log_e(\text{BMI}_{it}) = \beta_0 + \beta_1 \text{Age}_{it} + \sum_{k=1}^K \beta_{k+1} (\text{Age}_{it} - \kappa_k)_+ + b_{0i} + b_1 \text{Age}_{it} + \sum_{k=1}^K b_{(k+1)i} (\text{Age}_{it} - \kappa_k)_+ + \varepsilon_{it} \quad (2)$$

3. *LMM with cubic smoothing splines:*

$$\log_e(\text{BMI}_{it}) = \beta_0 + \sum_{j=1}^3 \beta_j \text{Age}_{it}^j + \sum_{k=1}^K \beta_{k+3} (\text{Age}_{it} - \kappa_k)_+^3 + b_{0i} + \sum_{j=1}^3 b_{ji} \text{Age}_{it}^j + \sum_{k=1}^K b_{(k+3)i} (\text{Age}_{it} - \kappa_k)_+^3 + \varepsilon_{it} \quad (3)$$

Where the fixed parameters are represented by $\beta_0, \beta_1, \beta_2, \dots, \kappa_k$ is the k^{th} knot and:

$$(t - \kappa_k)_+ = \begin{cases} 0 & \text{if } t \leq \kappa_k \\ t - \kappa_k & \text{if } t > \kappa_k \end{cases}$$

The parameter estimates for the random effects are represented by $b_{0i}, b_{1i}, b_{2i}, \dots$, and are assumed to be multivariate normally distributed. ε_{it} are the error terms assumed to be normally distributed and independent of the parameter estimates for the random effects. For the LMM with linear smoothing splines, the placement of knot points was based on the underlying biology of the BMI trajectory, with piecewise slopes for the time period of infancy to the AR (first knot point at 5.5

years), a slope from the AR to the pubertal period (second knot point at 11 years), and finally from the pubertal period through adolescence. We attempted fitting a third knot point at 9 months, but the model failed to converge. For the LMM with cubic smoothing splines, we *a priori* fit knot points at 2, 8 and 12 years based on previous studies (Warrington et al. 2013). For the LMMs with the cubic slope and cubic smoothing splines, we also compared the model fit when reducing the degree of the polynomial in the random effects to decrease the complexity of the model and attempt to assist in model convergence. For example, for the cubic slope LMM, we fit random effects that had a quadratic function of age and a linear function of age, in addition to the cubic function of age. A list of these models is given in the Supplementary Table 3.

The three LMMs were conducted on the $\log_e(\text{BMI})$ data within each cohort stratified for sex, and we conducted models with and without specifying a continuous autocorrelation structure of order 1 (CAR(1)) correlation structure.

Selection of the best model

Assessment of model fit was appraised using the following indices of model quality and goodness of fit; R^2 (conditional and marginal), intraclass correlation coefficient (ICC), Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), root mean squared error (RMSE), and residual SD. The selection of the overall best model was based on the most favourable model performance metrics across all cohorts, as well as model convergence and warning and error messages. It is important to note that the 'best fitting model' for each individual cohort was not necessarily chosen, as each cohort was slightly different for which model fit the data best.

The best model was taken forwards for knot placement refinement. We used a systematic approach to refine the best fitting model further using an incremental series of knot points where knot 1 was placed at 1, 1.5, and 2 years; knot 2 was placed at 6, 7, and 8 years and knot 3 was placed at 12 years for all models. As for the model diagnostics above, indices of model quality and goodness of fit were used to determine our final BMI trajectory model.

Estimating phenotypes for GWAS from the BMI trajectories

From the BMI trajectories generated by our model, we sought to extract interpretable characteristics that we could subsequently interrogate for genetic associations. We estimated the age and BMI at the AP and AR, two features of the BMI trajectory that respectively mark the transitions from the rise in BMI across infancy to the decline in BMI in early childhood, then the

subsequent rise in BMI from early childhood. In addition, we hypothesised that there may be shared and/or distinct genetic contributions to BMI preceding and succeeding these features (Couto Alves et al. 2019). To this end we defined time periods of approximately linear change in BMI prior to the AP, between the AP and AR, from the AR to the approximate onset of puberty and through adolescence. In addition to the rate of change (slopes) over these time intervals, we also analysed the area under the BMI curve (AUC), an estimate of the child's cumulative exposure to excessive body weight. We refer to each of these estimated parameters from the BMI trajectories as the estimated phenotypes.

Deriving the adiposity peak and rebound

We determined both the age (in years) and BMI (kg/m^2) at the AP and AR from the BMI trajectory model. For each participant, $\log_e(\text{BMI})$ was predicted using the fixed and random coefficients from the models at intervals of 0.01 year. Then, the AP was defined as the first maximal $\log_e(\text{BMI})$ occurring between the ages of 0.25 and 10 years of age, and the AR was determined as the first nadir within that time interval. These cut-off points were chosen based on previous evidence of the mean ages of AP and AR in European populations (Sovio et al. 2011). The age and BMI for these two points was estimated by taking the exponential of predicted $\log_e(\text{BMI})$. We ensured that the age at AP is less than the age at AR; when this did not occur, the individual's age and BMI at both AP and AR were set to missing.

Choosing the time intervals for deriving slopes and AUCs

A series of time periods of approximately linear change in BMI were defined, avoiding the inflection points at the AP and AR. The time intervals to derive these slopes (and AUCs) were determined visually from the population average BMI trajectories in each cohort and by minimising the proportion of individuals with AP and/or AR falling within the defined time periods (Supplementary Table 1). We chose the following time intervals for derivation of the slopes and AUCs:

- 0 to 6 months (referred to as infancy)
- 1 ½ years to 3 ½ years (early childhood)
- 6 ½ years to 10 years (late childhood)
- 12 years to 17 years (adolescence)

Although these time intervals avoided the average age of the AP and AR, there was a small proportion of individuals with AP and/or AR falling within these time intervals.

Deriving the slopes

Slopes for each individual for each time interval were estimated using:

$$Slope_{b-a} = \frac{y_b - y_a}{x_b - x_a}$$

Where y_b is the predicted value of log(BMI) from the LMM at the ending point for the time interval (e.g., 6 months for the infancy slope), y_a is the predicted value of log(BMI) from the LMM at the starting point (e.g., 2 weeks for the infancy slope), and x_a and x_b are the ages at the earlier and later time points, respectively.

Deriving the area under the curve (AUC)

The AUCs for each time interval were estimated by integrating the best fitting model with respect to age. For example, the AUC for the model with cubic smoothing splines from equation (3) above would be estimated using the following:

$$\int_a^b \left(\beta_0 + \sum_{j=1}^3 \beta_j Age_{it}^j + \sum_{k=1}^K \beta_{k+3} (Age_{it} - \kappa_k)_+^3 + b_{0i} + \sum_{j=1}^3 b_{ji} Age_{it}^j + \sum_{k=1}^K b_{(k+3)i} (Age_{it} - \kappa_k)_+^3 + \varepsilon_{it} \right) d(Age)$$

Where a is the earlier time point (e.g. 0 years for the infancy AUC) in each time interval and b is the later time point (e.g. 6 months for the infancy AUC).

Final model checks

Individuals were flagged as being an outlier for any of the estimated phenotypes (slopes, AUCs, age and BMI at AP/AR) based on the interquartile range (IQR) (Tukey JW 1977), with values outside twice the IQR flagged as outliers (*i.e.*, two times the IQR above the third quartile and below the first quartile). If an individual was flagged as an outlier for any one of the estimated phenotypes, they were excluded from analyses involving any of the estimated phenotypes. We took this conservative approach to excluding outliers, because if an individual was flagged as an outlier for an estimated phenotype it likely indicates an issue with their whole predicted curve. Per cohort, this excluded the following proportions of individuals from further analysis: 5% in NFBC1986, 6% in NFBC1966, 9% in ALSPAC, 10% CHOP African American, 12% OBE and 16% in CHOP European American sample.

Correlation matrices between the estimated phenotypes were generated to aid in downstream interpretation of GWAS results. We also conducted association analyses between each of the estimated phenotypes and BMI at the end of the trajectory for each cohort (*i.e.* BMI at age 16

years in OBE and 18 years for all other cohorts). We converted each of the estimated phenotypes and BMI at the end of the trajectory to z-scores by subtracting the mean and dividing by the standard deviation within each cohort so that we could compare across phenotypes. Analyses were adjusted for sex, and adjusted R^2 were reported.

GWAS of the estimated phenotypes

Genotyping in each of the contributing cohorts was performed using high-density Illumina BeadChip arrays, and data cleaning and quality control (QC) were performed locally for each cohort (see Supplementary Table 1 for details). Imputation for all European cohorts was performed using the reference data from the Haplotype Reference Consortium (HRC) release 1.1 (McCarthy et al. 2016). SNP positions were based on National Center for Biotechnology Information (NCBI) build 37 (hg19), and alleles were labelled on the positive strand of the reference genome. For the CHOP African American subset, imputation was performed using the TOPMED reference data (Kowalski et al. 2019) and SNP positions were based on NCBI build 38.

We performed sex-combined GWAS on the 12 estimated phenotypes derived from the BMI trajectory models: the four estimated slopes, the four AUCs, and age and BMI at AP and AR (OBE performed GWAS on 10 estimated phenotypes as they did not estimate the adolescent slope and AUC due to lack of data after 16 years). Ancestry principal components were included in the models as covariates, along with sex and cohort-specific covariates where appropriate. GWAS was performed using the imputed allelic dosage data under an additive genetic model. For ALSPAC, NFBC1966, NFBC1986, and OBE, a linear model was applied using PLINK 2.0 (Chang et al. 2015). For CHOP, analyses using a linear mixed model in REGENIE (version 3.2.6) was used to account for the extended family structure (Mbatchou et al. 2021)

Meta-analysis of the estimated phenotypes

Prior to meta-analysis, variants were first filtered at the cohort level for monomorphic or multiallelic SNPs, indels, low sample size (<20), low minor allele count (≤ 3), large effect estimates (absolute value of beta or SE ≥ 10 units per allele [log BMI units per year for the estimated slopes, log BMI years for the estimated AUCs, years for the estimated age at AP and AR and kg/m^2 for the estimated BMI at AP and AR]), and poor imputation quality (INFO <0.4 or R^2 <0.3). Filtering as well as harmonisation of variants across the cohorts and comparisons of allele frequencies was performed using the EasyQC2 (v. 1.1.1.b5) software package (Winkler et al. 2014) We conducted an inverse-variance weighted fixed-effects meta-analysis, combining each of the five European

cohorts, for each of the estimated phenotypes using GWAMA v.2.1 (Mägi and Morris 2010). OBE was excluded from the meta-analysis of the adolescent phenotypes as it did not have data after 16 years of age. Additionally, as OBE is an obesity cohort, we conducted a sensitivity meta-analysis excluding OBE from the other meta-analyses to ensure that it was not driving any observed association. After meta-analysis, we excluded variants that were not present in $\geq 50\%$ cohorts (i.e., 3 for the European cohorts; 2 for the European cohorts in the sensitivity analyses excluding OBE) as well as variants with minor allele frequency (MAF) < 0.005 . Results were clumped and annotated using LocusZoom ((Boughton et al. 2021), version 0.14.0). In the African American subset of CHOP, we selected any genome-wide significant variants identified in the European meta-analysis and performed lift over to build 37 using UCSC Lift Genome Annotations browser tool (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>).

Using the genome-wide summary statistics from the meta-analysis, we calculated the SNP-based heritability of each of the estimated phenotypes and the genetic correlations between phenotypes using LD score regression ((B. K. Bulik-Sullivan et al. 2015; B. Bulik-Sullivan et al. 2015), version 1.0.1). We also estimated genetic correlations between the derived phenotypes and cross-sectional BMI at birth, six months, 1.5 years, 3 years, 7 years, 8 years and adulthood using publicly available summary statistics. The specific ages for childhood BMI were selected so that they would be as close as possible to the cutoff points of our time windows of infancy, early and late childhood. The childhood BMI summary statistics from the Norwegian Mother, Father and Child Cohort Study (MoBA) were reported by (Helgeland et al. 2022). Summary statistics for adult BMI were obtained from the largest meta-analysis to-date of GWAS on cross-sectional adult BMI in individuals of European ancestry (Yengo et al. 2018). The summary statistics were formatted using a modified version of the supplied `munge.sumstats` python script. Variants with minor allele frequency less than 1% were removed; no filtering on imputation quality was performed as this was not available in the meta-analysis. SNP-based heritabilities and genetic correlations were calculated using the European LD scores available from the developers of LD score regression via the genomicSEM website (<https://utexas.box.com/s/vkd36n197m8klbaio3yzoxsee6sxo11v>).

Acknowledgments and funding:

Cohort acknowledgements and funding can be found in the Supplementary Text.

KB, DAL, and KT contributions to this work are supported by the UK Medical Research Council (MC_UU_00032/02 and MC_UU_00032/05). DAL's contribution is further supported by the British

Heart Foundation (CH/F/20/90003 and AA/18/1/34219) and the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement No 101021566). AH is supported by the European Union's Horizon 2020 research and innovation program under grant agreement No. 874739 (LongITools), and the Academy of Finland decision 336449 (Profi6). IP and ZB were funded in part by the Diabetes UK (British Diabetic Association no. 20/0006307), European Union's Horizon 2020 research and innovation program (H2020 Science with and for Society) (LONGITOOLS, H2020-SC1-2019-874739). IP is in part supported by BHF Basic Science Research fellowships (FS/15/59/31839 & FS/SBSRF/21/31025). SFAG is supported by the NICHD (R01 HD056465) and the Daniel B. Burke Endowed Chair for Diabetes Research. NMW is funded by an Australian National Health and Medical Research Council Investigator grant (APP2008723).

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Conflict of interest:

DAL received support from Medtronic Ltd and Roche Diagnostics for research unrelated to that presented here. All other authors report no conflict of interest.

References

- Boughton, Andrew P, Ryan P Welch, Matthew Flickinger, Peter VandeHaar, Daniel Taliun, Gonçalo R Abecasis, and Michael Boehnke. 2021. 'LocusZoom.js: Interactive and Embeddable Visualization of Genetic Association Study Results'. *Bioinformatics* 37 (18): 3017–18. <https://doi.org/10.1093/bioinformatics/btab186>.
- Boyd, Andy, Jean Golding, John Macleod, Debbie A. Lawlor, Abigail Fraser, John Henderson, Lynn Molloy, Andy Ness, Susan Ring, and George Davey Smith. 2013. 'Cohort Profile: The "Children of the 90s"—the Index Offspring of the Avon Longitudinal Study of Parents and Children'. *International Journal of Epidemiology* 42 (1): 111–27. <https://doi.org/10.1093/ije/dys064>.
- Brand, Judith S., Romy Gaillard, Jane West, Rosemary R. C. McEachan, John Wright, Ellis Voerman, Janine F. Felix, Kate Tilling, and Deborah A. Lawlor. 2019. 'Associations of Maternal Quitting, Reducing, and Continuing Smoking during Pregnancy with Longitudinal Fetal Growth: Findings from Mendelian Randomization and Parental Negative Control Studies'. Edited by Cosetta Minelli. *PLOS Medicine* 16 (11): e1002972. <https://doi.org/10.1371/journal.pmed.1002972>.
- Bulik-Sullivan, Brendan, Hilary K. Finucane, Verner Anttila, Alexander Gusev, Felix R. Day, Po-Ru Loh, ReproGen Consortium, et al. 2015. 'An Atlas of Genetic Correlations across Human Diseases and Traits'. *Nature Genetics* 47 (11): 1236–41. <https://doi.org/10.1038/ng.3406>.
- Bulik-Sullivan, Brendan K., Po-Ru Loh, Hilary K. Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J. Daly, Alkes L. Price, and Benjamin M. Neale. 2015. 'LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies'. *Nature Genetics* 47 (3): 291–95. <https://doi.org/10.1038/ng.3211>.
- Campbell, Michele Wen-Chien, Joanne Williams, John B. Carlin, and Melissa Wake. 2011. 'Is the Adiposity Rebound a Rebound in Adiposity?' *International Journal of Pediatric Obesity: IJPO: An Official Journal of the International Association for the Study of Obesity* 6 (2–2): e207-215. <https://doi.org/10.3109/17477166.2010.526613>.
- Canouil, Mickaël, Nicole Warrington, Kimberley Burrows, and Anni Heiskala. 2024. 'Eggl: Early Growth Genetics Longitudinal Analysis'. Zenodo. <https://doi.org/10.5281/zenodo.10594717>.
- Chang, Christopher C, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, and James J Lee. 2015. 'Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets'. *GigaScience* 4 (1): s13742-015-0047-0048. <https://doi.org/10.1186/s13742-015-0047-8>.
- Cole, T. J. 2004. 'Children Grow and Horses Race: Is the Adiposity Rebound a Critical Period for Later Obesity?' *BMC Pediatrics* 4 (March): 6. <https://doi.org/10.1186/1471-2431-4-6>.
- Cole, T. J., J. V. Freeman, and M. A. Preece. 1995. 'Body Mass Index Reference Curves for the UK, 1990'. *Archives of Disease in Childhood* 73 (1): 25–29. <https://doi.org/10.1136/adc.73.1.25>.
- Connolly, John J, Joseph T Glessner, Dong Li, Patrick MA Sleiman, and Hakon Hakonarson. 2020. 'The Center for Applied Genomics at The Children's Hospital of Philadelphia – Pediatric Perspectives on Genomic Medicine'. *The Journal of Precision Medicine* (blog). 2020. <https://www.thejournalofprecisionmedicine.com/the-journal-of-precision-medicine/the-center-for-applied-genomics-at-the-childrens-hospital-of-philadelphia-pediatric-perspectives-on-genomic-medicine/>.
- Couto Alves, Alexsander, N. Maneka G. De Silva, Ville Karhunen, Ulla Sovio, Shikta Das, H. Rob Taal, Nicole M. Warrington, et al. 2019. 'GWAS on Longitudinal Growth Traits

- Reveals Different Genetic Factors Influencing Infant, Child, and Adult BMI'. *Science Advances* 5 (9): eaaw3095. <https://doi.org/10.1126/sciadv.aaw3095>.
- Daymont, Carrie, Michelle E Ross, A Russell Localio, Alexander G Fiks, Richard C Wasserman, and Robert W Grundmeier. 2017. 'Automated Identification of Implausible Values in Growth Data from Pediatric Electronic Health Records'. *Journal of the American Medical Informatics Association* 24 (6): 1080–87. <https://doi.org/10.1093/jamia/ocx037>.
- DiTomasso, Diane, and Mary Cloud. 2019. 'Systematic Review of Expected Weight Changes After Birth for Full-Term, Breastfed Newborns'. *Journal of Obstetric, Gynecologic & Neonatal Nursing* 48 (6): 593–603. <https://doi.org/10.1016/j.jogn.2019.09.004>.
- Elhakeem, Ahmed, Monika Frysz, Ana G. Soares, Joshua A. Bell, Tim J. Cole, Jon Heron, Laura D. Howe, et al. 2023. 'Evaluation and Comparison of Nine Growth- and Development-Based Measures of Pubertal Timing'. medRxiv. <https://doi.org/10.1101/2023.06.12.23290796>.
- Elks, Cathy E., Ruth J. F. Loos, Stephen J. Sharp, Claudia Langenberg, Susan M. Ring, Nicholas J. Timpson, Andrew R. Ness, et al. 2010. 'Genetic Markers of Adult Obesity Risk Are Associated with Greater Early Infancy Weight Gain and Growth'. *PLoS Medicine* 7 (5): e1000284. <https://doi.org/10.1371/journal.pmed.1000284>.
- Fraser, Abigail, Corrie Macdonald-Wallis, Kate Tilling, Andy Boyd, Jean Golding, George Davey Smith, John Henderson, et al. 2013. 'Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC Mothers Cohort'. *International Journal of Epidemiology* 42 (1): 97–110. <https://doi.org/10.1093/ije/dys066>.
- Goldstein, Harvey, and Bianca De Stavola. 2010. 'Statistical Modelling of Repeated Measurement Data'. *Longitudinal and Life Course Studies* 1 (2): 170–85. <https://doi.org/10.14301/lcs.v1i2.67>.
- Gouveia, Mateus H., Amy R. Bentley, Hampton Leonard, Karlijn A. C. Meeks, Kenneth Ekoru, Guanjie Chen, Michael A. Nalls, et al. 2021. 'Trans-Ethnic Meta-Analysis Identifies New Loci Associated with Longitudinal Blood Pressure Traits'. *Scientific Reports* 11 (February): 4075. <https://doi.org/10.1038/s41598-021-83450-3>.
- Gouveia, Mateus H., Cibele C. Cesar, Meddy L. Santolalla, Hanaisa P. Sant Anna, Marilia O. Scliar, Thiago P. Leal, Nathalia M. Araújo, et al. 2019. 'Genetics of Cognitive Trajectory in Brazilians: 15 Years of Follow-up from the Bambuí-Epigen Cohort Study of Aging'. *Scientific Reports* 9 (1): 18085. <https://doi.org/10.1038/s41598-019-53988-4>.
- Hardy, Rebecca, Andrew K. Wills, Andrew Wong, Cathy E. Elks, Nicholas J. Wareham, Ruth J. F. Loos, Diana Kuh, and Ken K. Ong. 2010. 'Life Course Variations in the Associations between FTO and MC4R Gene Variants and Body Size'. *Human Molecular Genetics* 19 (3): 545–52. <https://doi.org/10.1093/hmg/ddp504>.
- Helgeland, Øyvind, Marc Vaudel, Pol Sole-Navais, Christopher Flatley, Julius Juodakis, Jonas Bacelis, Ingvild L. Koløen, et al. 2022. 'Characterization of the Genetic Architecture of Infant and Early Childhood Body Mass Index'. *Nature Metabolism* 4 (3): 344–58. <https://doi.org/10.1038/s42255-022-00549-1>.
- Howe, Laura D., Priyakumari G. Parmar, Lavinia Paternoster, Nicole M. Warrington, John P. Kemp, Laurent Briollais, John P. Newnham, et al. 2013. 'Genetic Influences on Trajectories of Systolic Blood Pressure across Childhood and Adolescence'. *Circulation. Cardiovascular Genetics* 6 (6): 608–14. <https://doi.org/10.1161/CIRCGENETICS.113.000197>.
- Järvelin, M. R., P. Elliott, I. Kleinschmidt, M. Martuzzi, C. Grundy, A. L. Hartikainen, and P. Rantakallio. 1997. 'Ecological and Individual Predictors of Birthweight in a Northern Finland Birth Cohort 1986'. *Paediatric and Perinatal Epidemiology* 11 (3): 298–312. <https://doi.org/10.1111/j.1365-3016.1997.tb00007.x>.
- Ko, Seyoon, Christopher A. German, Aubrey Jensen, Judong Shen, Anran Wang, Devan V. Mehrotra, Yan V. Sun, Janet S. Sinsheimer, Hua Zhou, and Jin J. Zhou. 2022. 'GWAS of

- Longitudinal Trajectories at Biobank Scale'. *American Journal of Human Genetics* 109 (3): 433–45. <https://doi.org/10.1016/j.ajhg.2022.01.018>.
- Kowalski, Madeline H., Huijun Qian, Ziyi Hou, Jonathan D. Rosen, Amanda L. Tapia, Yue Shan, Deepti Jain, et al. 2019. 'Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium Whole Genome Sequences Improves Imputation Quality and Detection of Rare Variant Associations in Admixed African and Hispanic/Latino Populations'. *PLoS Genetics* 15 (12): e1008500. <https://doi.org/10.1371/journal.pgen.1008500>.
- Laird, N. M., and J. H. Ware. 1982. 'Random-Effects Models for Longitudinal Data'. *Biometrics* 38 (4): 963–74.
- Mägi, Reedik, and Andrew P. Morris. 2010. 'GWAMA: Software for Genome-Wide Association Meta-Analysis'. *BMC Bioinformatics* 11 (1): 288. <https://doi.org/10.1186/1471-2105-11-288>.
- Mbatchou, Joelle, Leland Barnard, Joshua Backman, Anthony Marcketta, Jack A. Kosmicki, Andrey Ziyatdinov, Christian Benner, et al. 2021. 'Computationally Efficient Whole-Genome Regression for Quantitative and Binary Traits'. *Nature Genetics* 53 (7): 1097–1103. <https://doi.org/10.1038/s41588-021-00870-7>.
- McCarthy, Shane, Sayantan Das, Warren Kretzschmar, Olivier Delaneau, Andrew R. Wood, Alexander Teumer, Hyun Min Kang, et al. 2016. 'A Reference Panel of 64,976 Haplotypes for Genotype Imputation'. *Nature Genetics* 48 (10): 1279–83. <https://doi.org/10.1038/ng.3643>.
- Meirelles, Osorio D., Jun Ding, Toshiko Tanaka, Serena Sanna, Hsih-Te Yang, Dawood B. Dudekula, Francesco Cucca, Luigi Ferrucci, Goncalo Abecasis, and David Schlessinger. 2013. 'SHAVE: Shrinkage Estimator Measured for Multiple Visits Increases Power in GWAS of Quantitative Traits'. *European Journal of Human Genetics: EJHG* 21 (6): 673–79. <https://doi.org/10.1038/ejhg.2012.215>.
- Meyre, David, Jérôme Delplanque, Jean-Claude Chèvre, Cécile Lecoeur, Stéphane Lobbens, Sophie Gallina, Emmanuelle Durand, et al. 2009. 'Genome-Wide Association Study for Early-Onset and Morbid Adult Obesity Identifies Three New Risk Loci in European Populations'. *Nature Genetics* 41 (2): 157–59. <https://doi.org/10.1038/ng.301>.
- Meyre, David, Cécile Lecoeur, Jérôme Delplanque, Stephan Francke, Vincent Vatin, Emmanuelle Durand, Jacques Weill, Christian Dina, and Philippe Froguel. 2004. 'A Genome-Wide Scan for Childhood Obesity-Associated Traits in French Families Shows Significant Linkage on Chromosome 6q22.31-Q23.2'. *Diabetes* 53 (3): 803–11. <https://doi.org/10.2337/diabetes.53.3.803>.
- Min, Jungwon, Xiaozhong Wen, Hong Xue, and Youfa Wang. 2018. 'Ethnic Disparities in Childhood BMI Trajectories and Obesity and Potential Causes among 29,250 US Children: Findings from the Early Childhood Longitudinal Study-Birth and Kindergarten Cohorts'. *International Journal of Obesity* 42 (9): 1661–70. <https://doi.org/10.1038/s41366-018-0091-4>.
- Molenberghs, Geert, and Geert Verbeke. 2000. *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. New York, NY: Springer. <https://doi.org/10.1007/978-1-4419-0300-6>.
- Mook-Kanamori, D. O., L. Ay, A. Hofman, C. M. van Duijn, H. A. Moll, H. Raat, A. C. S. Hokken-Koelega, and V. W. V. Jaddoe. 2011. 'No Association of Obesity Gene FTO with Body Composition at the Age of 6 Months. The Generation R Study'. *Journal of Endocrinological Investigation* 34 (1): 16–20. <https://doi.org/10.3275/7075>.
- Nordström, Tanja, Jouko Miettunen, Juha Auvinen, Leena Ala-Mursula, Sirkka Keinänen-Kiukaanniemi, Juha Veijola, Marjo-Riitta Järvelin, Sylvain Sebert, and Minna Männikkö. 2021. 'Cohort Profile: 46 Years of Follow-up of the Northern Finland Birth Cohort 1966

- (NFBC1966). *International Journal of Epidemiology* 50 (6): 1786–87. <https://doi.org/10.1093/ije/dyab109>.
- Plachta-Danielzik, S., A. Bosy-Westphal, B. Kehden, M. I. Gehrke, K. Kromeyer-Hauschild, M. Grillenberger, C. Willhöft, S. B. Heymsfield, and M. J. Müller. 2013. 'Adiposity Rebound Is Misclassified by BMI Rebound'. *European Journal of Clinical Nutrition* 67 (9): 984–89. <https://doi.org/10.1038/ejcn.2013.131>.
- R Core Team. 2022. 'R: A Language and Environment for Statistical Computing.' R. Vienna, Austria.: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Robinson, G. K. 1991. 'That BLUP Is a Good Thing: The Estimation of Random Effects'. *Statistical Science* 6 (1): 15–32. <https://doi.org/10.1214/ss/1177011926>.
- Rolland-Cachera, M. F., M. Deheeger, F. Bellisle, M. Sempé, M. Guilloud-Bataille, and E. Patois. 1984. 'Adiposity Rebound in Children: A Simple Indicator for Predicting Obesity'. *The American Journal of Clinical Nutrition* 39 (1): 129–35. <https://doi.org/10.1093/ajcn/39.1.129>.
- Sikorska, Karolina, Emmanuel Lesaffre, Patrick J. F. Groenen, Fernando Rivadeneira, and Paul H. C. Eilers. 2018. 'Genome-Wide Analysis of Large-Scale Longitudinal Outcomes Using Penalization —GALLOP Algorithm'. *Scientific Reports* 8 (1): 6815. <https://doi.org/10.1038/s41598-018-24578-7>.
- Sikorska, Karolina, Fernando Rivadeneira, Patrick J. F. Groenen, Albert Hofman, André G. Uitterlinden, Paul H. C. Eilers, and Emmanuel Lesaffre. 2013. 'Fast Linear Mixed Model Computations for Genome-Wide Association Studies with Longitudinal Data'. *Statistics in Medicine* 32 (1): 165–80. <https://doi.org/10.1002/sim.5517>.
- Silventoinen, Karri, Weiling Li, Aline Jelenkovic, Reijo Sund, Yoshie Yokoyama, Sari Aaltonen, Maarit Piirtola, et al. 2022. 'Changing Genetic Architecture of Body Mass Index from Infancy to Early Adulthood: An Individual Based Pooled Analysis of 25 Twin Cohorts'. *International Journal of Obesity (2005)* 46 (10): 1901–9. <https://doi.org/10.1038/s41366-022-01202-3>.
- Smith, Andrew Dac, Rebecca Hardy, Jon Heron, Carol J. Joinson, Debbie A. Lawlor, Corrie Macdonald-Wallis, and Kate Tilling. 2016. 'A Structured Approach to Hypotheses Involving Continuous Exposures over the Life Course'. *International Journal of Epidemiology* 45 (4): 1271–79. <https://doi.org/10.1093/ije/dyw164>.
- Smith, Erin N., Wei Chen, Mika Kähönen, Johannes Kettunen, Terho Lehtimäki, Leena Peltonen, Olli T. Raitakari, et al. 2010. 'Longitudinal Genome-Wide Association of Cardiovascular Disease Risk Factors in the Bogalusa Heart Study'. *PLOS Genetics* 6 (9): e1001094. <https://doi.org/10.1371/journal.pgen.1001094>.
- Sovio, Ulla, Dennis O. Mook-Kanamori, Nicole M. Warrington, Robert Lawrence, Laurent Briollais, Colin N. A. Palmer, Joanne Cecil, et al. 2011. 'Association between Common Variation at the FTO Locus and Changes in Body Mass Index from Infancy to Late Childhood: The Complex Nature of Genetic Association through Growth and Development'. *PLoS Genetics* 7 (2): e1001307. <https://doi.org/10.1371/journal.pgen.1001307>.
- Stergiakouli, Evangelia, Romy Gaillard, Jeremy M Tavaré, Nina Balthasar, Ruth J Loos, Hendrik R Taal, David M Evans, et al. 2014. 'Genome-Wide Association Study of Height-Adjusted BMI in Childhood Identifies Functional Variant in ADCY3'. *Obesity (Silver Spring, Md.)* 22 (10): 2252–59. <https://doi.org/10.1002/oby.20840>.
- Thorleifsson, Gudmar, G. Bragi Walters, Daniel F. Gudbjartsson, Valgerdur Steinthorsdottir, Patrick Sulem, Anna Helgadóttir, Unnur Styrkarsdóttir, et al. 2009. 'Genome-Wide Association Yields New Sequence Variants at Seven Loci That Associate with Measures of Obesity'. *Nature Genetics* 41 (1): 18–24. <https://doi.org/10.1038/ng.274>.
- Tukey JW. 1977. *Exploratory Data Analysis*. https://scholar.google.com/scholar_lookup?&title=Exploratory%20Data%20Analysis&publication_year=1977&author=Tukey%2CJ.W.

- Turley, Patrick, Raymond K. Walters, Omeed Maghizian, Aysu Okbay, James J. Lee, Mark Alan Fontana, Tuan Anh Nguyen-Viet, et al. 2018. 'Multi-Trait Analysis of Genome-Wide Association Summary Statistics Using MTAG'. *Nature Genetics* 50 (2): 229–37. <https://doi.org/10.1038/s41588-017-0009-4>.
- Venkatesh, Samvida S., Habib Ganjgahi, Duncan S. Palmer, Kayesha Coley, Laura B. L. Wittemans, Christoffer Nellaker, Chris Holmes, Cecilia M. Lindgren, and George Nicholson. 2023. 'The Genetic Architecture of Changes in Adiposity during Adulthood'. medRxiv. <https://doi.org/10.1101/2023.01.09.23284364>.
- Warrington, Nicole M., Laura D. Howe, Lavinia Paternoster, Marika Kaakinen, Sauli Herrala, Ville Huikari, Yan Yan Wu, et al. 2015. 'A Genome-Wide Association Study of Body Mass Index across Early Life and Childhood'. *International Journal of Epidemiology* 44 (2): 700–712. <https://doi.org/10.1093/ije/dyv077>.
- Warrington, Nicole M., Laura D. Howe, Yan Yan Wu, Nicholas J. Timpson, Kate Tilling, Craig E. Pennell, John Newnham, et al. 2013. 'Association of a Body Mass Index Genetic Risk Score with Growth throughout Childhood and Adolescence'. *PLoS ONE* 8 (11): e79547. <https://doi.org/10.1371/journal.pone.0079547>.
- Wen, Xiaozhong, Ken Kleinman, Matthew W. Gillman, Sheryl L. Rifas-Shiman, and Elsie M. Taveras. 2012. 'Childhood Body Mass Index Trajectories: Modeling, Characterizing, Pairwise Correlations and Socio-Demographic Predictors of Trajectory Characteristics'. *BMC Medical Research Methodology* 12 (1): 38. <https://doi.org/10.1186/1471-2288-12-38>.
- Wendel, Bernadette, Sergi Papiol, Till F. M. Andlauer, Jörg Zimmermann, Jens Wiltfang, Carsten Spitzer, Fanny Senner, et al. 2021. 'A Genome-Wide Association Study of the Longitudinal Course of Executive Functions'. *Translational Psychiatry* 11 (1): 1–8. <https://doi.org/10.1038/s41398-021-01510-8>.
- Winkler, Thomas W., Felix R. Day, Damien C. Croteau-Chonka, Andrew R. Wood, Adam E. Locke, Reedik Mägi, Teresa Ferreira, et al. 2014. 'Quality Control and Conduct of Genome-Wide Association Meta-Analyses'. *Nature Protocols* 9 (5): 1192–1212. <https://doi.org/10.1038/nprot.2014.071>.
- Yengo, Loic, Julia Sidorenko, Kathryn E. Kemper, Zhili Zheng, Andrew R. Wood, Michael N. Weedon, Timothy M. Frayling, et al. 2018. 'Meta-Analysis of Genome-Wide Association Studies for Height and Body Mass Index in ~700000 Individuals of European Ancestry'. *Human Molecular Genetics* 27 (20): 3641–49. <https://doi.org/10.1093/hmg/ddy271>.
- Yuan, Min, Xu Steven Xu, Yaning Yang, Yinsheng Zhou, Yi Li, Jinfeng Xu, Jose Pinheiro, and Alzheimer's Disease Neuroimaging Initiative. 2021. 'SCEBE: An Efficient and Scalable Algorithm for Genome-Wide Association Studies on Longitudinal Outcomes with Mixed-Effects Modeling'. *Briefings in Bioinformatics* 22 (3): bbaa130. <https://doi.org/10.1093/bib/bbaa130>.

Figure 1: Population average BMI trajectories predicted by our final chosen model (cubic spline function in the fixed effects with cubic slope function in the random effects) for each of the six cohorts for males (A) and females (B). BMI trajectories were predicted from 2 weeks to 17 years, which corresponds to the age range that the slopes and AUCs were predicted from, in all cohorts except OBE where they were predicted from 2 weeks to 16 years due to the lack of data after age 16. The year(s) of recruitment for each cohort are as follows: ALSPAC: 1991-1993, CHOP: 1988-present, NFBC1966: 1966, NFBC1986: 1985-1986, OBE: 1981-2001

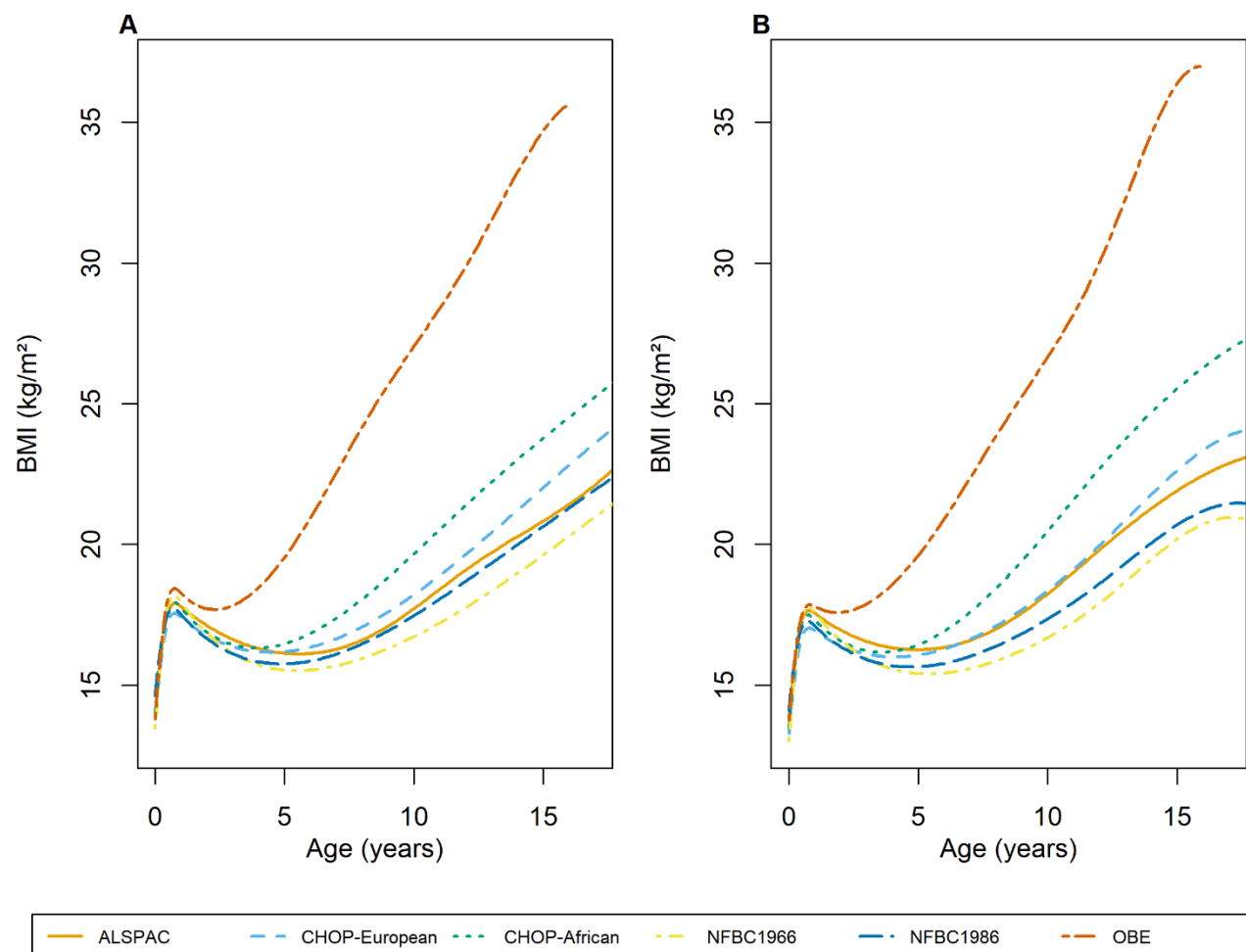


Figure 2: Observed BMI measures versus fitted BMI values from our final chosen model (cubic spline function in the fixed effects with cubic slope function in the random effects) in males. A) ALSPAC; B) CHOP European subset; C) CHOP African American subset; D) NFBC1966; E) NFBC1986; F) OBE. BMI is given in kg/m^2

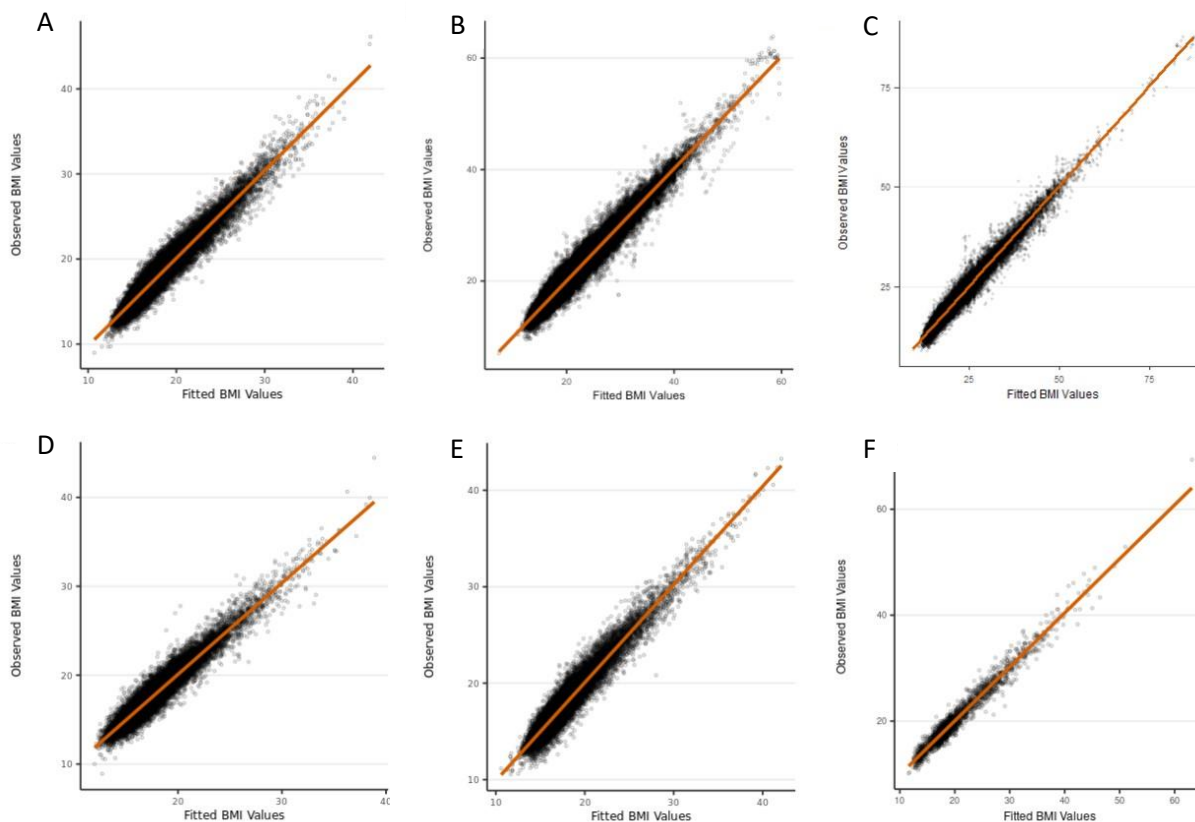


Figure 3: SNP-based heritability and genome-wide genetic correlation between the twelve estimated phenotypes summarising growth across early life. SNP-based heritabilities, presented on the diagonals, and genetic correlations, presented on the off diagonals, were derived using linkage disequilibrium score regression. SNP-based heritabilities for AP age, infancy and adolescent slope are low, with high standard errors, and therefore the genetic correlations with these traits are unreliable but are shown for completeness. *estimates of genetic correlation were >1; given this is not possible we have set these to one.

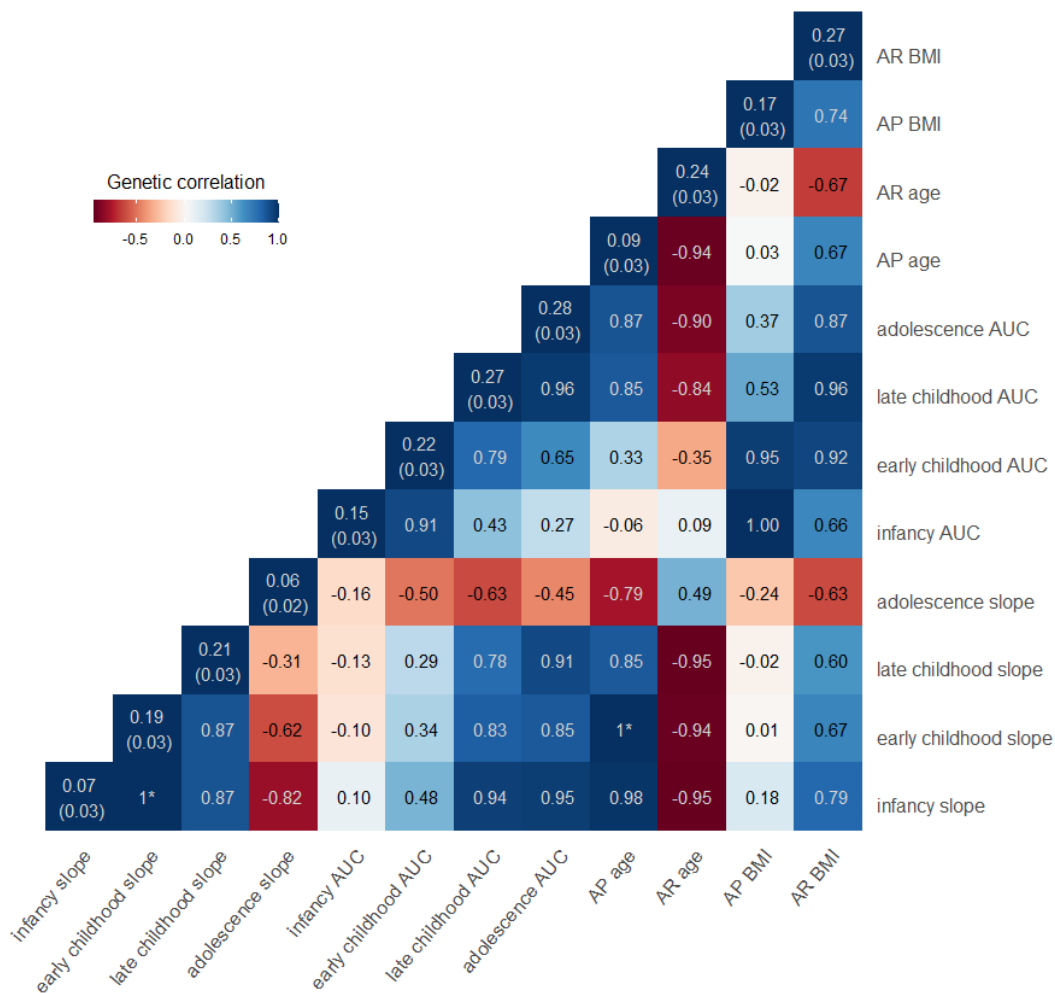


Figure 4: Manhattan plots and quantile-quantile (QQ) plots of the meta-analyses for the area under the curve estimated phenotypes across infancy (0-0.5 years), early childhood (1.5-3.5 years), late childhood (6.5-10 years) and adolescence (12-17 years). Loci are labelled with their nearest gene annotated by LocusZoom. The red dotted line corresponds to the genome-wide significance level of $P < 5 \times 10^{-8}$.

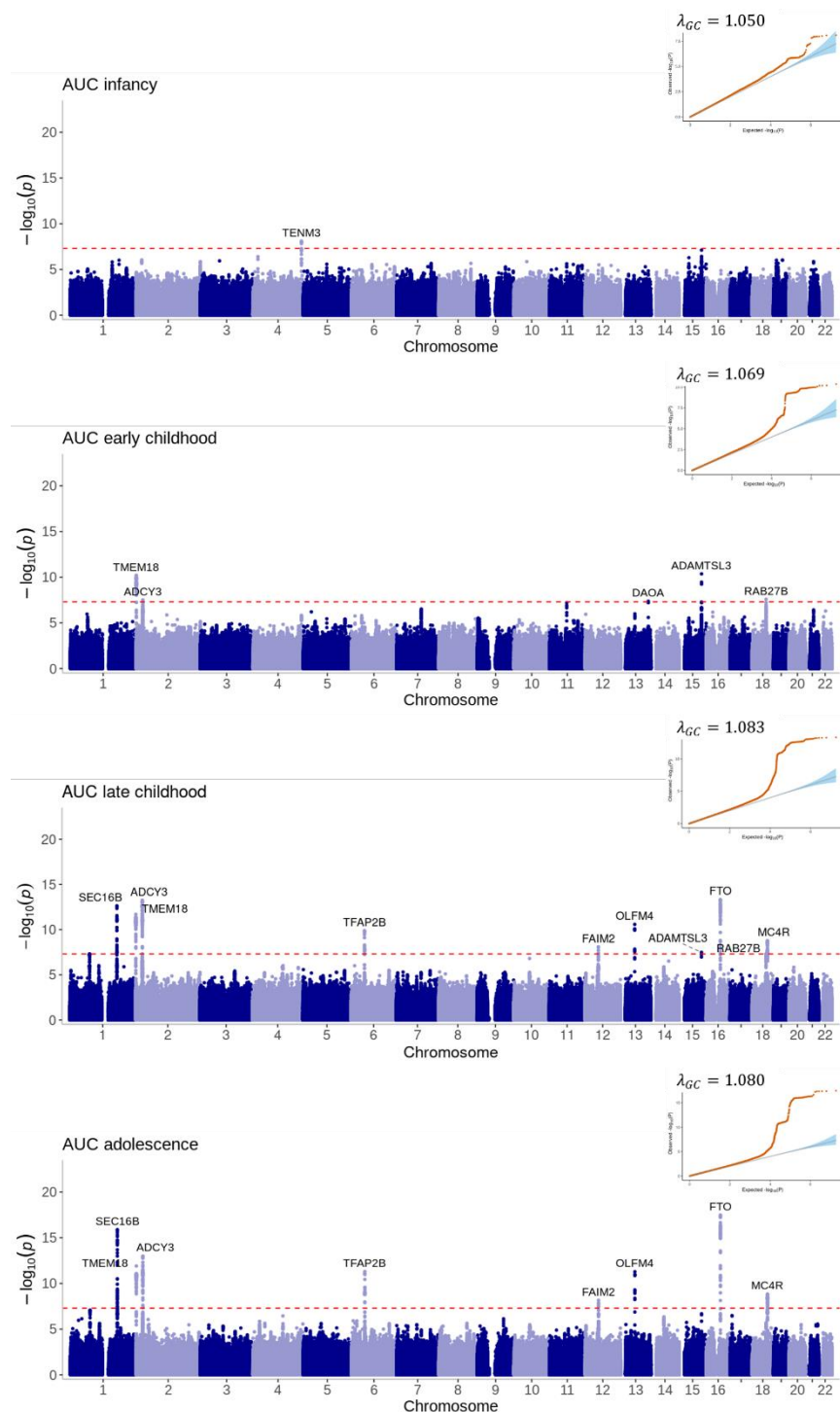


Figure 5: Manhattan plots and quantile-quantile (QQ) plots of the meta-analyses for the slope estimated phenotypes across infancy (0-0.5 years), early childhood (1.5-3.5 years), late childhood (6.5-10 years) and adolescence (12-17 years). Loci are labelled with their nearest gene annotated by LocusZoom. The red dotted line corresponds to the genome-wide significance level of $P < 5 \times 10^{-8}$.

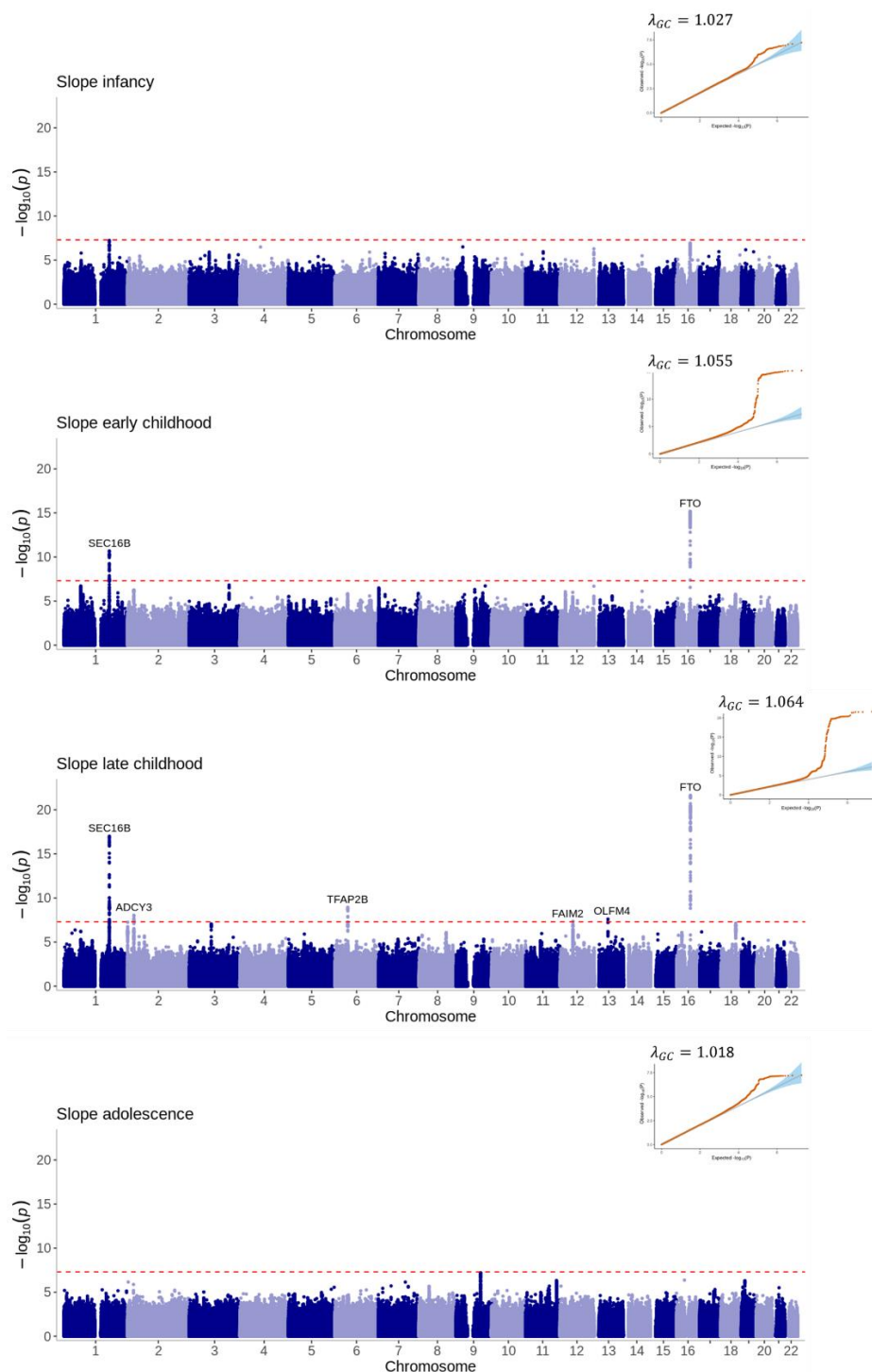


Figure 6: Manhattan plots and quantile-quantile (QQ) plots of the meta-analyses for the age and BMI at adiposity peak and adiposity rebound estimated phenotypes. Loci are labelled with their nearest gene annotated by LocusZoom. The red dotted line corresponds to the genome-wide significance level of $P < 5 \times 10^{-8}$.

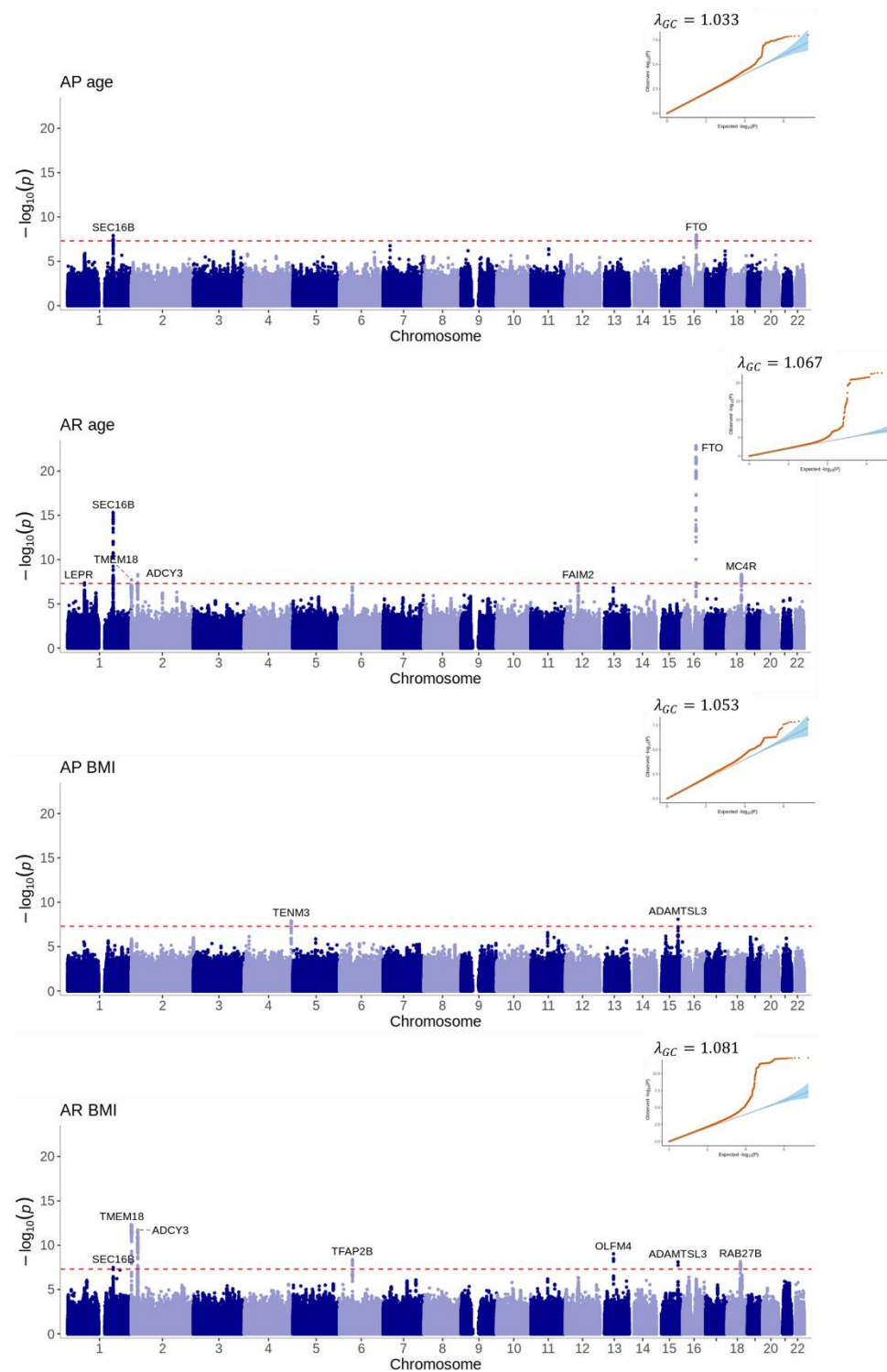


Table 1: Cohort information for cohort included in the analysis, including summary statistics of the estimated phenotypes.

See Supplementary Table 1 for further information. The summary statistics are presented as mean (standard deviation). The OBE cohort was unable to estimate the adolescent phenotypes due to lack of data after the age of 16 years (N/A for not available).

<i>Cohort information</i>	ALSPAC		CHOP - European American		CHOP - African American		NFBC1966		NFBC1986		OBE	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Country	Bristol, UK		Philadelphia, USA		Philadelphia, USA		Northern Finland		Northern Finland		France	
Ethnicity	European		European		African American		European		European		European	
Collection type	Pregnancy cohort		Random collection from Hospital visits		Random collection from Hospital visits		Prospective general population-based		Prospective general population-based		Retrospective obese children cohort	
Year of birth	1991-1993		1988-present		1988-present		1966		1985-1986		1981-2001	
Sample size	7197	6818	12002	10814	10533	10772	3800	3280	2734	2826	252	308
Number of BMI measures per person	8.48 (5.41)	8.83 (5.41)	13.90 (16.60)	13.90 (16.40)	13.40 (12.40)	12.60 (12.00)	15.64 (5.67)	15.90 (5.74)	19.31 (5.14)	19.54 (5.32)	9.59 (2.93)	9.59 (2.91)
Age at adiposity peak (yrs)	0.77 (0.04)	0.78 (0.04)	0.75 (0.04)	0.77 (0.05)	0.72 (0.04)	0.74 (0.05)	0.73 (0.03)	0.75 (0.04)	0.70 (0.03)	0.73 (0.04)	0.75 (0.08)	0.79 (0.06)
BMI at adiposity peak (kg/m ²)	17.75 (1.00)	17.33 (0.99)	17.58 (0.82)	17.02 (0.92)	17.94 (1.02)	17.46 (1.08)	18.22 (1.15)	17.81 (1.17)	17.75 (1.09)	17.29 (1.09)	18.39 (1.03)	17.83 (1.03)
Age at adiposity rebound (yrs)	5.69 (1.19)	5.10 (1.34)	4.90 (1.51)	4.54 (1.54)	4.22 (1.20)	3.83 (1.20)	5.67 (0.99)	5.61 (1.20)	5.06 (1.01)	4.87 (1.21)	2.40 (0.67)	2.12 (0.62)
BMI at adiposity rebound (kg/m ²)	15.78 (1.09)	15.75 (1.17)	15.78 (1.06)	15.57 (1.13)	15.96 (1.27)	15.82 (1.43)	15.40 (1.00)	15.25 (1.16)	15.59 (1.08)	15.48 (1.17)	17.34 (1.40)	17.28 (1.19)
Derived infancy slope (0-0.5 years)	0.45 (0.02)	0.47 (0.02)	0.44 (0.02)	0.46 (0.02)	0.46 (0.02)	0.48 (0.02)	0.56 (0.02)	0.58 (0.02)	0.36 (0.01)	0.38 (0.01)	0.54 (0.05)	0.48 (0.03)
Derived childhood slope (1.5-3.5 years)	-0.03 (0.01)	-0.02 (0.01)	-0.03 (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.02 (0.02)	-0.04 (0.01)	-0.04 (0.01)	-0.03 (0.01)	-0.03 (0.01)	0.01 (0.03)	0.01 (0.02)
Derived late childhood slope (6.5-10 years)	0.02 (0.01)	0.03 (0.01)	0.03 (0.02)	0.03 (0.02)	0.04 (0.02)	0.05 (0.02)	0.02 (0.01)	0.02 (0.01)	0.03 (0.01)	0.02 (0.01)	0.06 (0.02)	0.06 (0.01)
Derived adolescent slope (12-17 years)	0.03 (0.01)	0.03 (0.01)	0.04 (0.02)	0.04 (0.02)	0.03 (0.02)	0.03 (0.02)	0.03 (0.01)	0.03 (0.01)	0.03 (0.02)	0.03 (0.02)	N/A	N/A
Derived infancy AUC (0-0.5 years)	1.39 (0.03)	1.38 (0.03)	1.38 (0.03)	1.36 (0.03)	1.39 (0.03)	1.38 (0.03)	1.39 (0.03)	1.37 (0.03)	1.40 (0.03)	1.38 (0.03)	1.40 (0.03)	1.38 (0.03)
Derived childhood AUC (1.5-3.5 years)	5.65 (0.11)	5.64 (0.12)	5.61 (0.11)	5.57 (0.14)	5.62 (0.14)	5.59 (0.16)	5.61 (0.12)	5.58 (0.13)	5.59 (0.12)	5.56 (0.13)	5.75 (0.19)	5.75 (0.16)
Derived late childhood AUC (6.5-10 years)	9.86 (0.32)	9.94 (0.37)	9.93 (0.46)	9.95 (0.50)	10.14 (0.57)	10.24 (0.66)	9.70 (0.28)	9.69 (0.34)	9.82 (0.34)	9.80 (0.39)	11.19 (0.51)	11.12 (0.44)
Derived adolescent AUC (12-17 years)	15.07 (0.57)	15.30 (0.66)	15.32 (0.96)	15.43 (0.94)	15.73 (1.21)	16.09 (1.28)	14.76 (0.51)	14.87 (0.54)	15.01 (0.65)	15.01 (0.64)	N/A	N/A

Table 2: Model fit parameters from each cohort for a selection of the methods. See Supplementary Table 3 for further information on the model parameters from the other methods.

Model details	Diagnostic	Cohort												
		ALSPAC		CHOP - European American		CHOP - African American		NFBC1966		NFBC1986		OBE		
		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	
Cubic slope / Cubic slope	R ² conditional	0.83	Warning: Singular precision matrix in level -1	0.96	0.96	0.88	0.92	0.80	0.81	0.84	0.84	0.84	0.86	
	R ² marginal	0.31		0.32	0.35	0.48	0.58	0.30	0.31	0.30	0.31	0.68	0.70	
	ICC	0.75		0.93	0.94	0.76	0.80	0.72	0.73	0.77	0.77	0.52	0.52	
	RMSE	0.06		0.04	0.04	0.07	0.07	0.05	0.06	0.05	0.05	0.10	0.09	
	Sigma	0.06		0.05	0.05	0.07	0.07	0.06	0.06	0.05	0.05	0.10	0.10	
Linear splines / Linear splines	R ² conditional	0.84	0.86	0.95	Warning: Singular precision matrix in level -1	0.94	0.96	Warning: Singular precision matrix in level -1	0.82	0.85	0.84	0.85	0.86	
	R ² marginal	0.33	0.40	0.33		0.35	0.44		0.32	0.31	0.32	0.68	0.70	
	ICC	0.75	0.78	0.93		0.91	0.92		0.73	0.78	0.77	0.53	0.51	
	RMSE	0.06	0.06	0.05		0.05	0.05		0.05	0.05	0.05	0.05	0.09	0.09
	Sigma	0.06	0.07	0.05		0.06	0.06		0.06	0.05	0.05	0.05	0.10	0.10
Cubic splines / Cubic splines	R ² conditional	0.91	0.92	Warning: Singular precision matrix in level -1	Warning: Singular precision matrix in level -1	0.91	0.94	Warning: Singular precision matrix in level -1	0.89	0.90	0.90	0.93	0.92	
	R ² marginal	0.37	0.42			0.62	0.72		0.37	0.34	0.35	0.73	0.74	
	ICC	0.85	0.87			0.75	0.77		0.83	0.85	0.85	0.73	0.69	
	RMSE	0.04	0.04			0.05	0.05		0.04	0.04	0.04	0.06	0.06	
	Sigma	0.05	0.05			0.06	0.06		0.05	0.04	0.04	0.07	0.07	
Cubic splines / Cubic slope	R ² conditional	0.85	0.87	0.94	0.95	0.91	0.94	0.86	0.86	0.86	0.83	0.91	0.90	
	R ² marginal	0.53	0.63	0.47	0.42	0.58	0.71	0.42	0.46	0.45	0.48	0.84	0.83	
	ICC	0.68	0.66	0.88	0.92	0.78	0.78	0.75	0.74	0.74	0.68	0.42	0.42	
	RMSE	0.04	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.07	0.07	
	Sigma	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.04	0.05	0.08	0.08	