

Supplementary figures and text for:

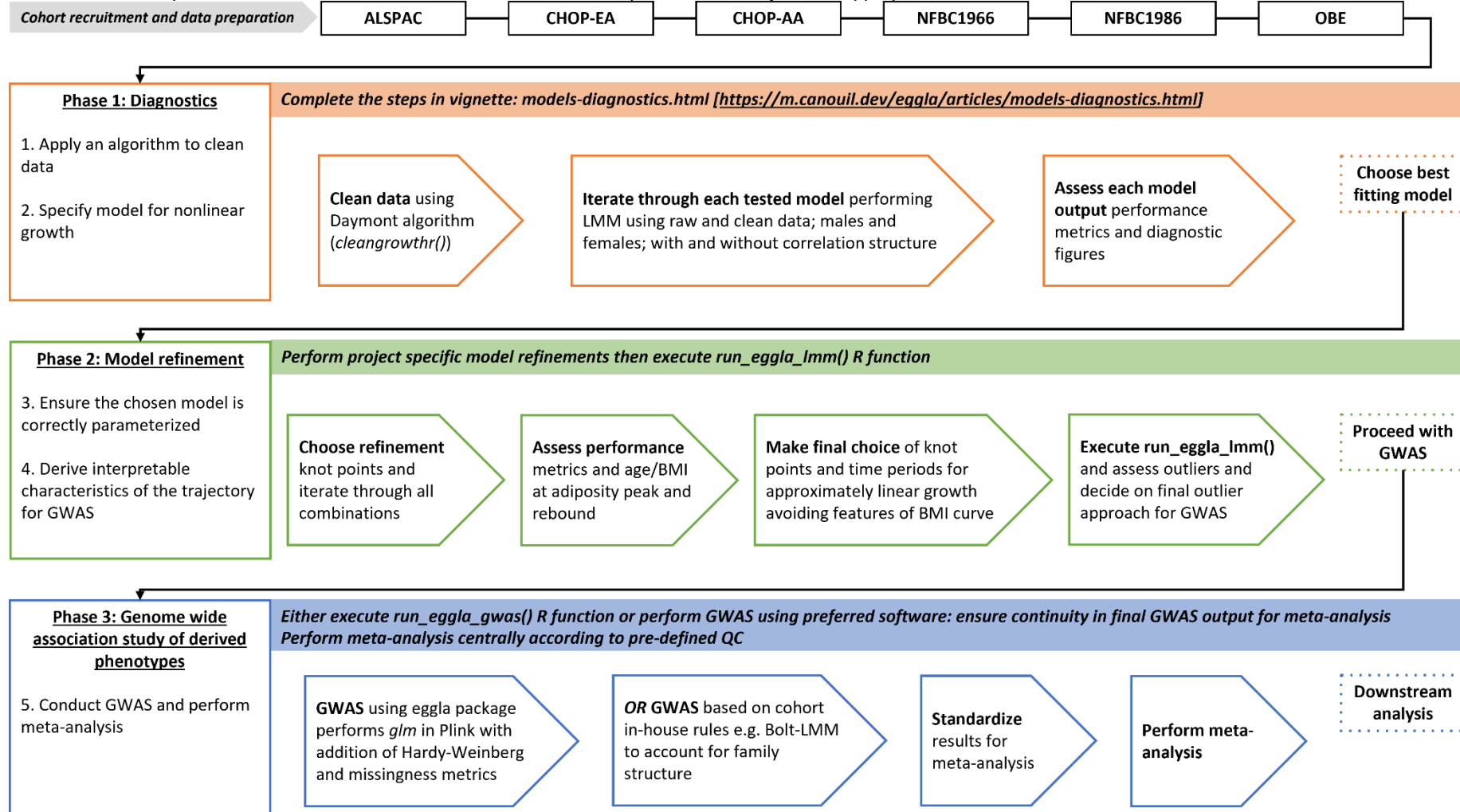
A framework for conducting time-varying genome-wide association studies: An application to body mass index across childhood in six multiethnic cohorts

Supplementary Figure 1: Flow diagram of analysis pipeline	2
Supplementary Figure 2: Refinement of knot points in the cubic spline with cubic slope random effects model	3
A: ALSPAC	3
B: CHOP-EA	4
C: CHOP-AA	5
D: NFBC1966	6
E: NFBC1986	7
F: OBE	8
Supplementary Figure 3: Residual plots for the chosen model	9
1: ALSPAC	9
2: CHOP-EA	11
3: CHOP-AA	13
4: NFBC1966	15
5: NFBC1986	17
6: OBE	19
Supplementary Figure 4: Summary of the association between final BMI (between age 16-18 years) and each of the estimated phenotypes, including a random effects meta-analysis (DerSimonian-Laird estimator)	21
Supplementary Figure 5: Manhattan plots of the meta-analyses without OBE for the 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 line corresponds to the genome-wide significance level of $P < 5 \times 10^{-8}$	24
Supplementary Figure 6: LocusZoom regional plot of the association between adolescent slope and the FAM120AOS locus	27
Cohort description, acknowledgements and funding	27
Avon Longitudinal Study of Parents and Children (ALSPAC)	27
Northern Finland Birth Cohorts born in 1966 (NFBC1966)	28
Northern Finland Birth Cohorts born in 1986 (NFBC1986)	29
CHOP	29
OBE	29
Linear mixed modelling	30
Example process and output from the <i>EGGLA</i> BMI modelling framework	31
References	33

Supplementary Figure 1: Flow diagram of analysis pipeline

The flow diagram illustrates the step by step procedure used to implement, test, assess and finalize non-linear BMI curve modeling ready for GWAS. Future analysts may consult this flow diagram and read further information in the Supplementary text about the outputs from the *EGGLA* pipeline.

ALSPAC Avon Longitudinal Study of Parents and Children; CHOP-EA Children's Hospital of Philadelphia European subset; CHOP-AA Children's Hospital of Philadelphia African American subset; NFBC1966 Northern Finland Birth Cohort 1966; NFBC1986 Northern Finland Birth Cohort 1986; LMM linear mixed model; GWAS genome-wide association study; BMI body mass index; glm generalized linear model; PCs principal components; QC quality control. *Here model refinement is specific to the present BMI growth analysis. For other growth methods, model refinement will be specific to the chosen model and knot refinements for spline models may not be appropriate.

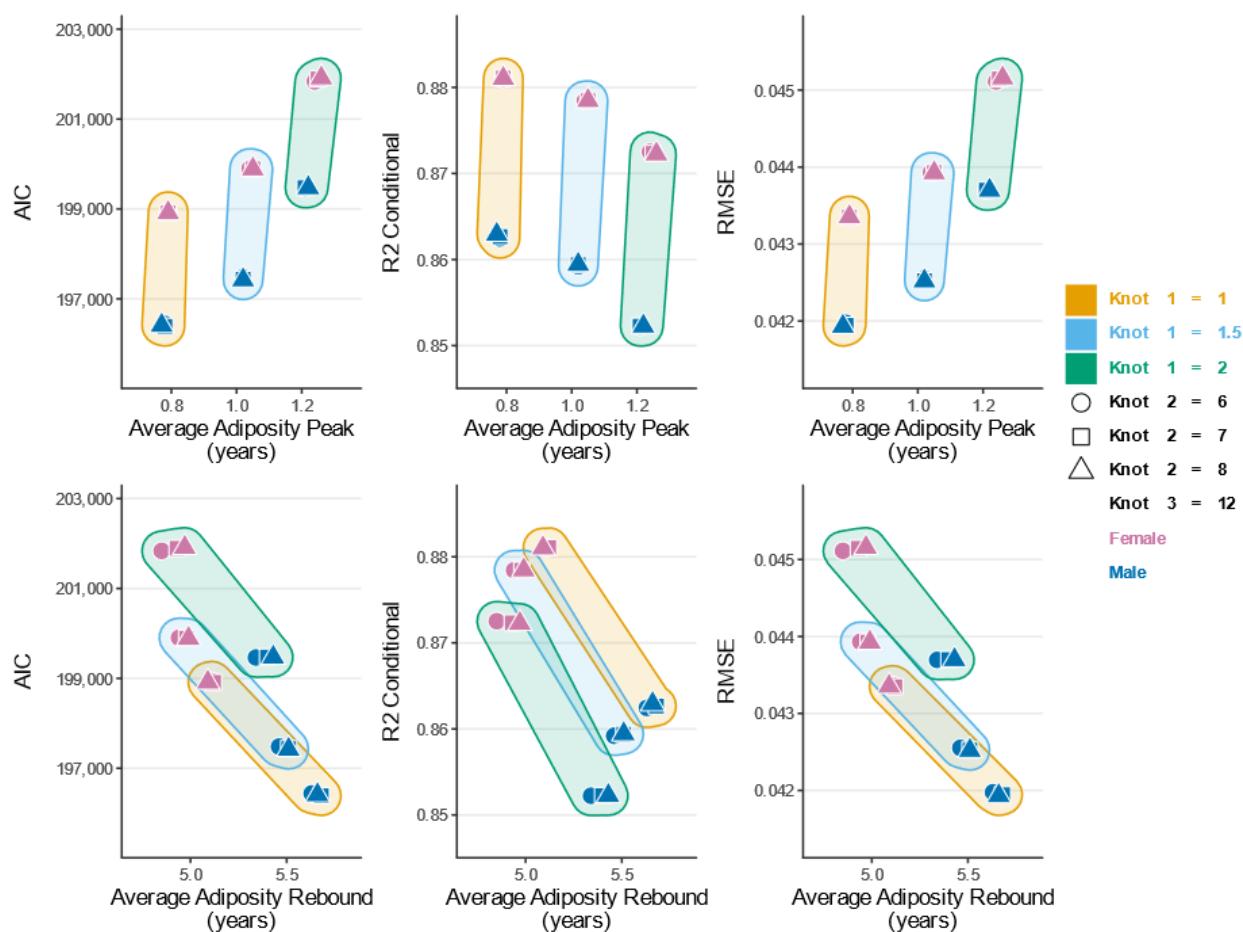


Supplementary Figure 2: Refinement of knot points in the cubic spline with cubic slope random effects model

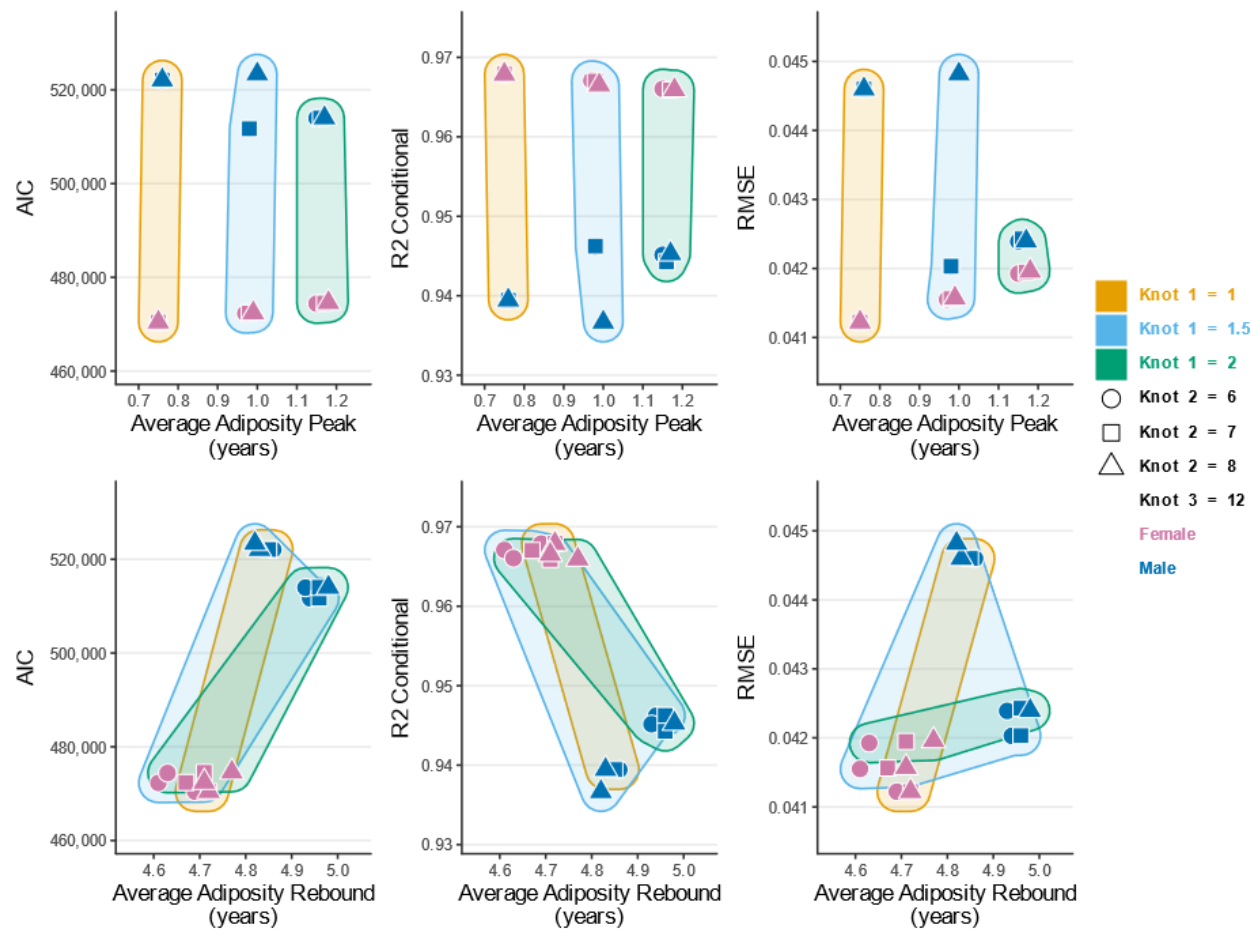
The cubic spline with cubic slope random effects models were selected for refinement. Models iterated pairwise through increments of knot 1 (1, 1.5 and 2 years) and knot 2 (6, 7, and 8 years), with knot 3 being held constant at 12 years. The figure displays the performance metrics AIC, conditional R^2 , and RMSE as a function of average age at adiposity peak and rebound. The knot point combinations are differentially coloured for knot 1 (yellow for 1 year, blue for 1.5 years, and green for 2 years) and has a differential marker symbol to denote knot 2 (circle for 6 years, square for 7 years, and triangle for 8 years). Marker colour indicates males (blue) and females (pink). Figures are given for ALSPAC [A], CHOP-EA [B], CHOP-AA [C], NFBC1966 [D], NFBC1986 [E], and OBE [F].

ALSPAC Avon Longitudinal Study of Parents and Children; CHOP-EA Children's Hospital of Philadelphia European subset; CHOP-AA Children's Hospital of Philadelphia African American subset; NFBC1966 Northern Finland Birth Cohort 1966; NFBC1986 Northern Finland Birth Cohort 1986; AIC Akaike information criterion ; RMSE root mean square error.

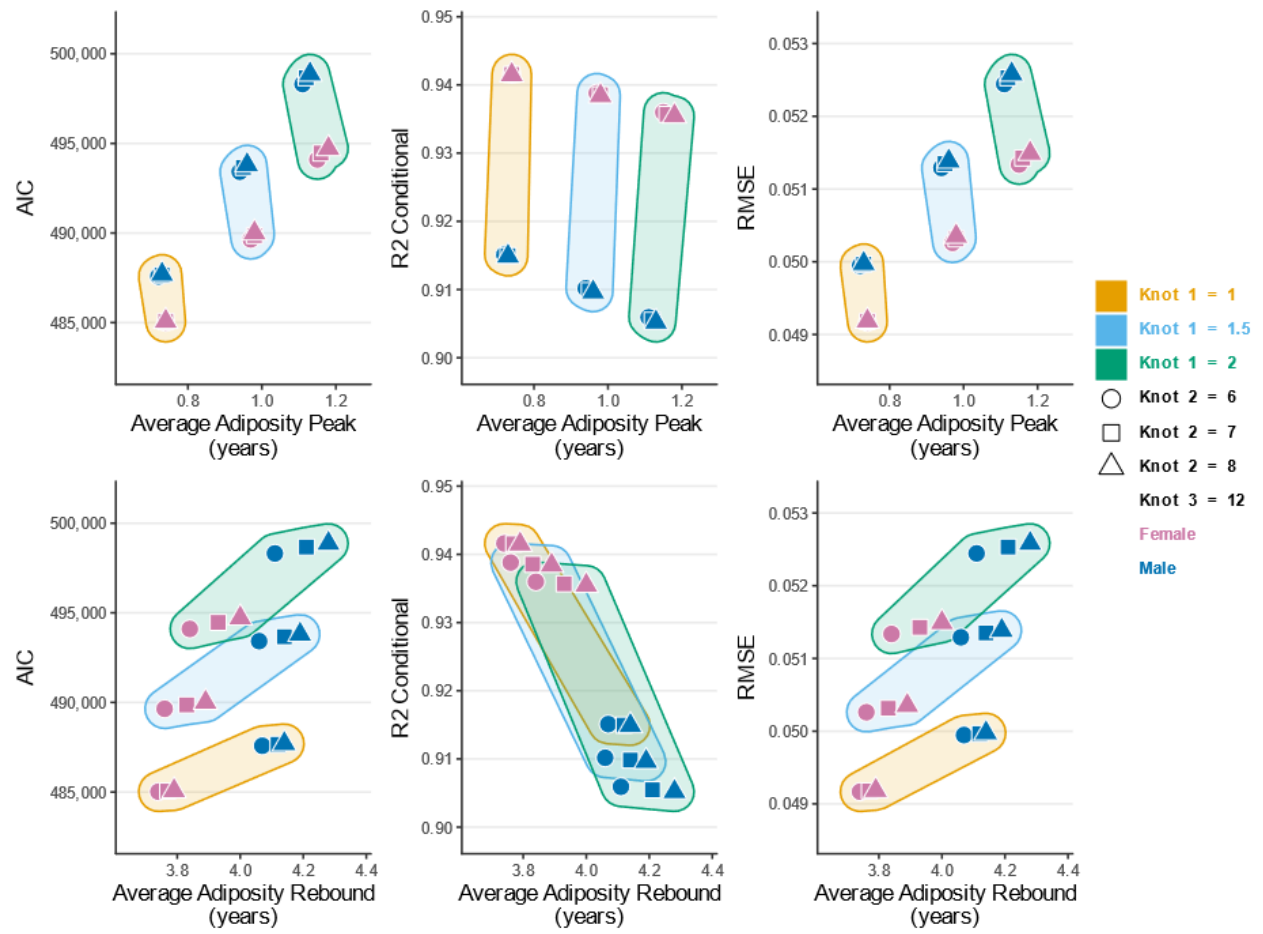
A: ALSPAC



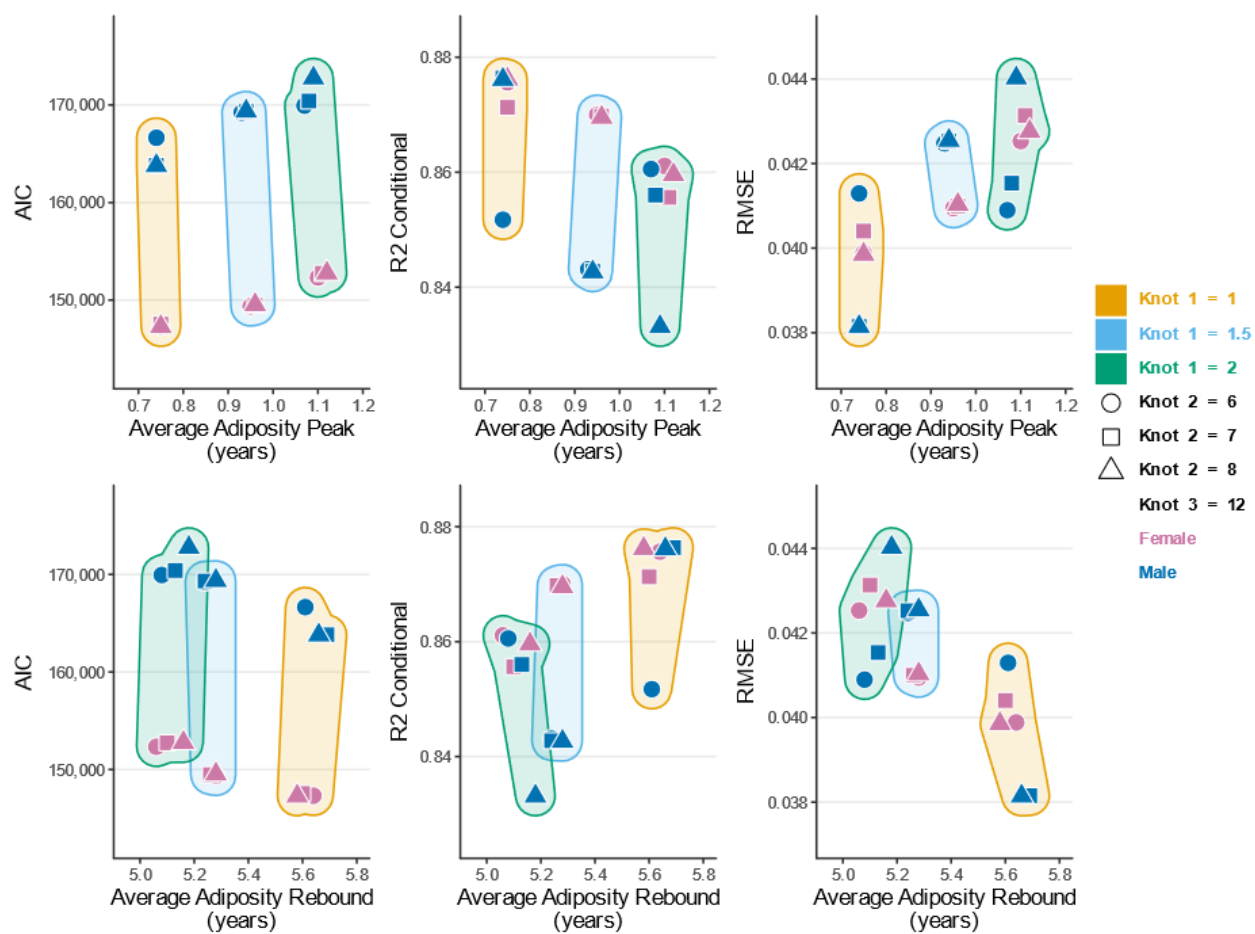
B: CHOP-EA



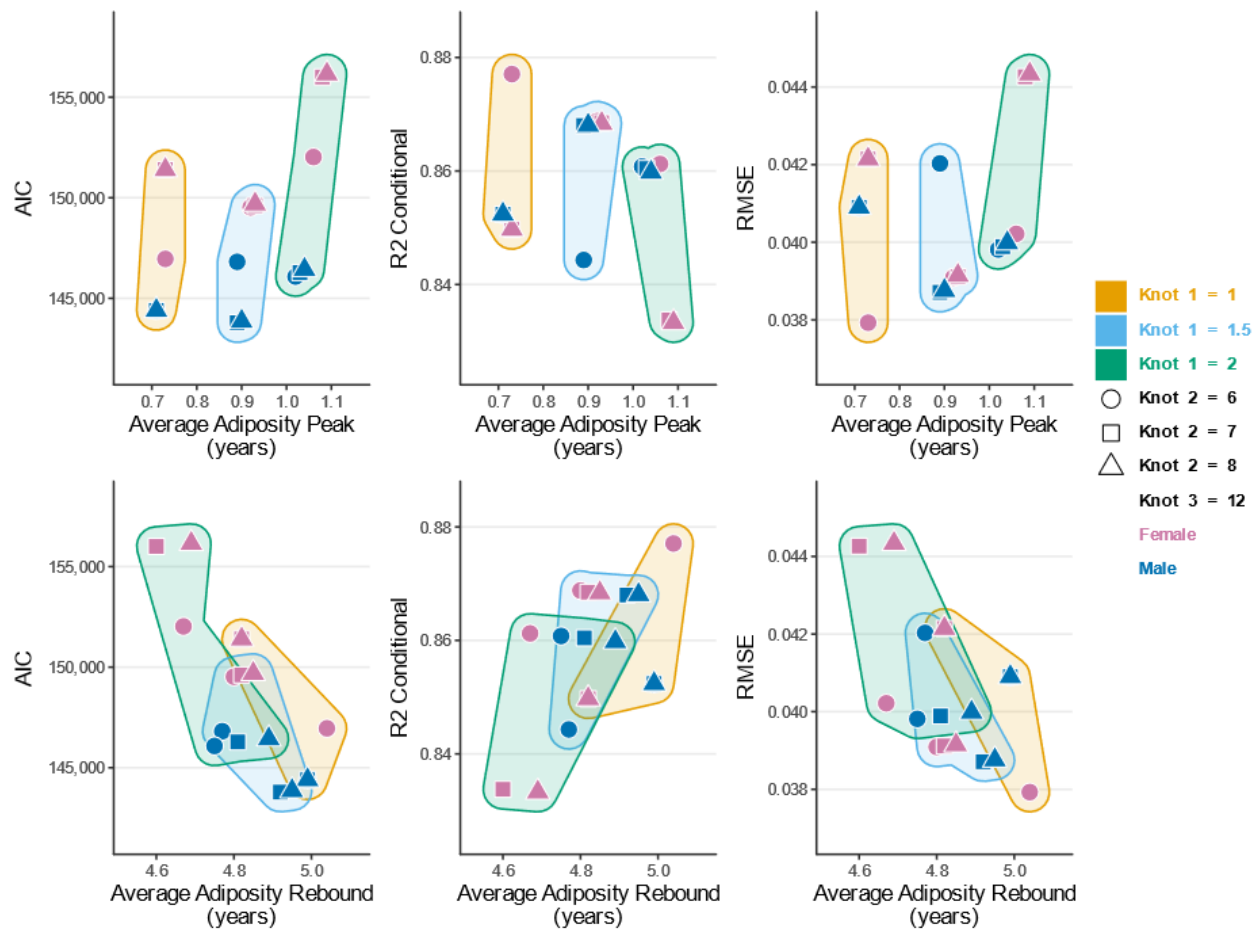
C: CHOP-AA



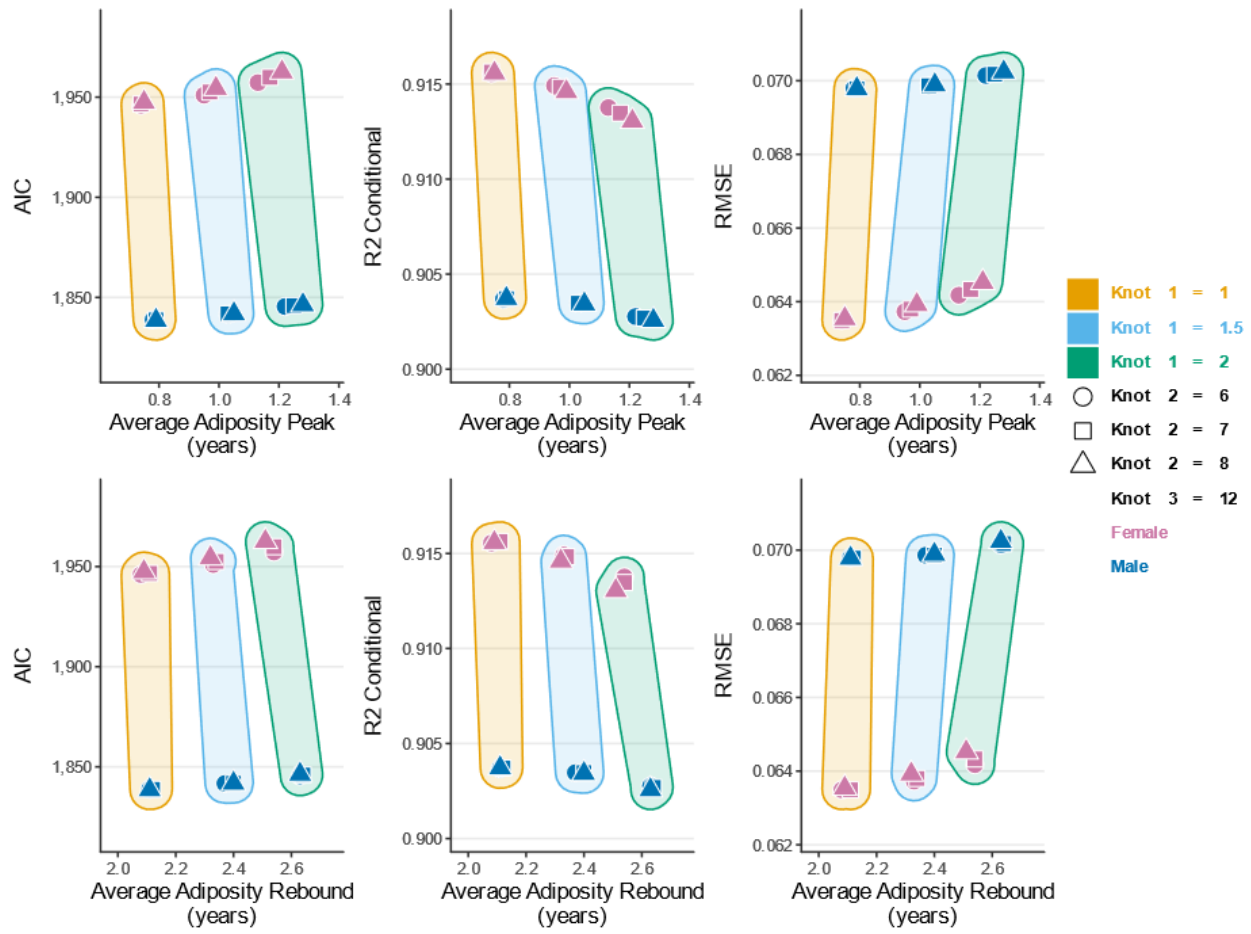
D: NFBC1966



E: NFBC1986



F: OBE



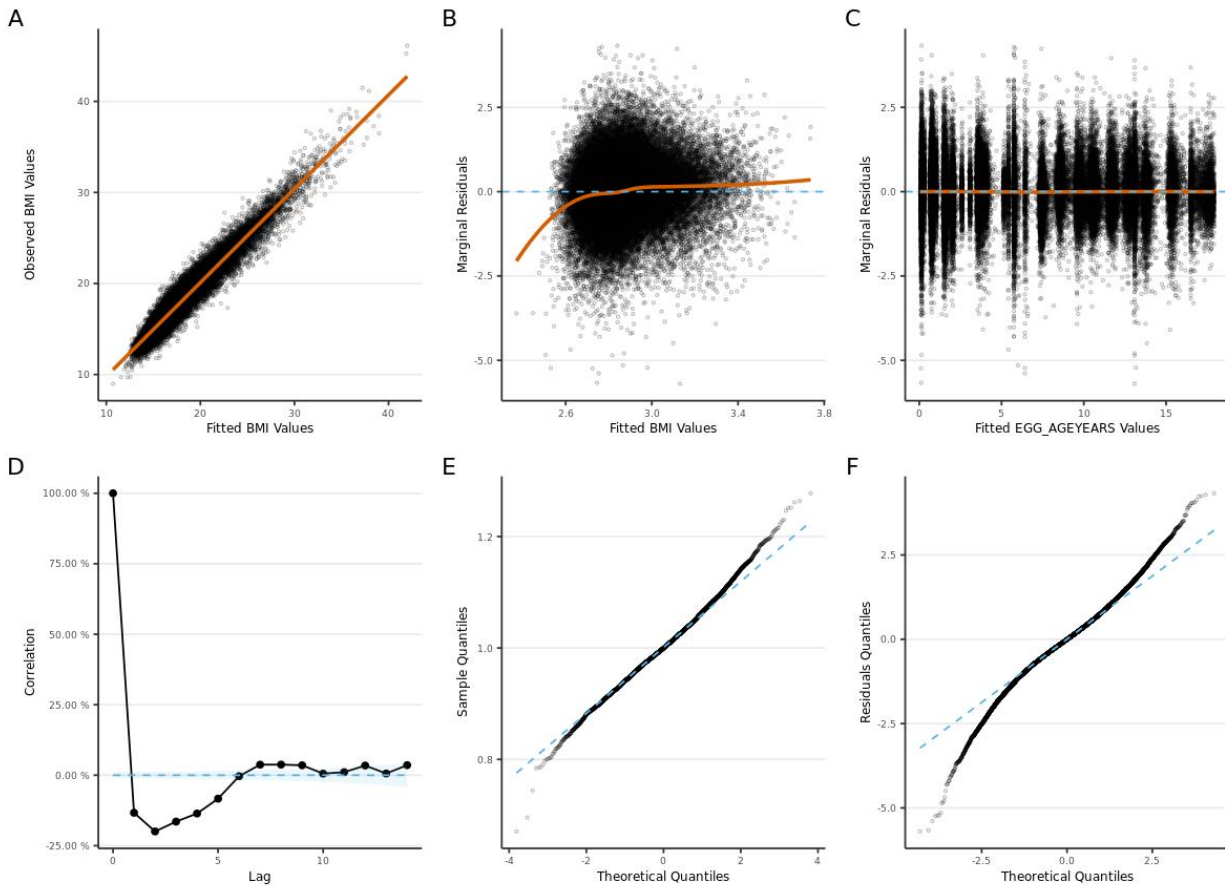
Supplementary Figure 3: Residual plots for the chosen model

Assessment of model fit for our final chosen model (cubic spline with cubic slope random effects). Each plot depicts: A) the fitted BMI values vs. observed B) the fitted BMI values vs. marginal residuals C) the fitted age values vs. marginal residuals D) autocorrelation function (ACF) vs. lag of normalized residuals E) theoretical vs. sample quantiles and F) theoretical vs. residual quantiles. 1) ALSPAC; 2) CHOP-EA; 3) CHOP-AA 4) NFBC1966; 5) NFBC1986; 6) OBE.

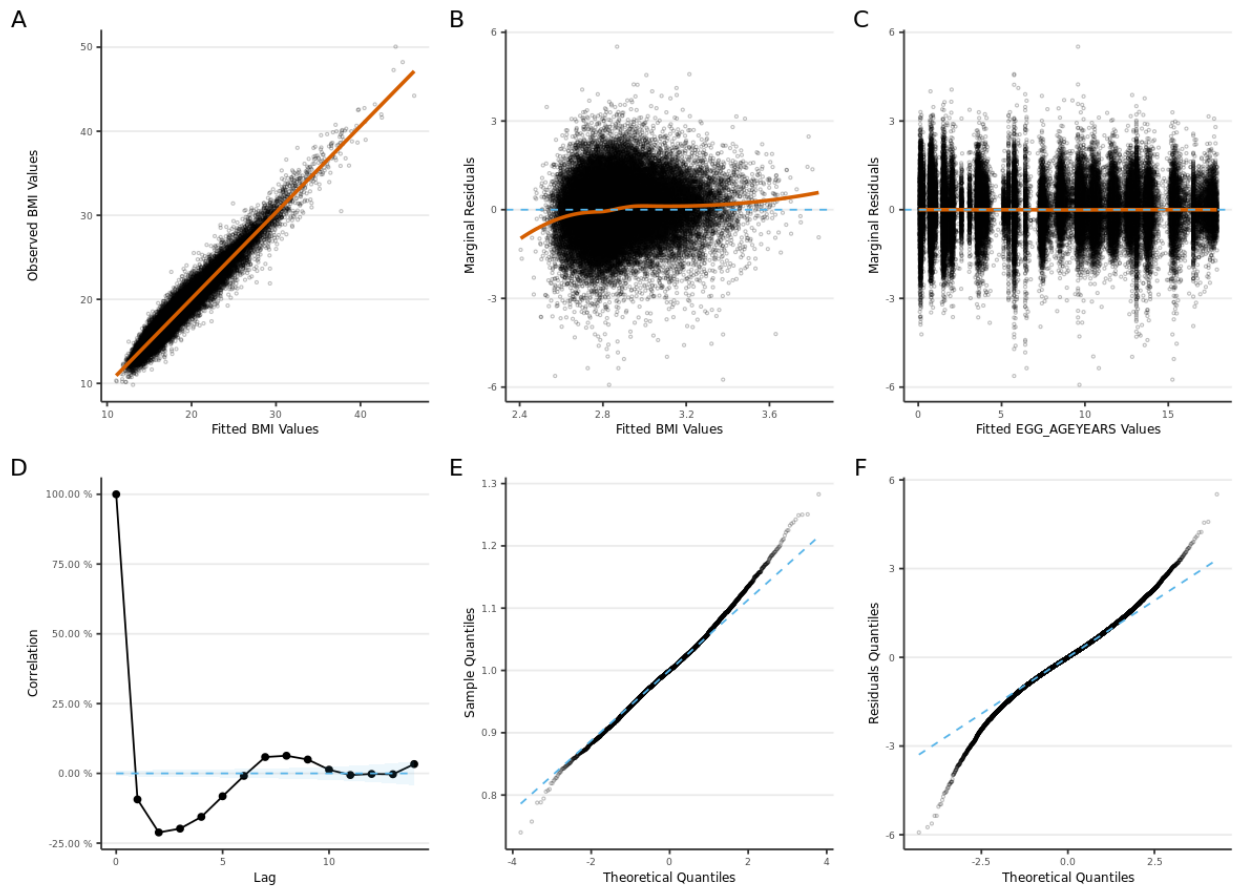
ALSPAC Avon Longitudinal Study of Parents and Children; CHOP-EA Children's Hospital of Philadelphia European subset; CHOP-AA Children's Hospital of Philadelphia African American subset; NFBC1966 Northern Finland Birth Cohort 1966; NFBC1986 Northern Finland Birth Cohort 1986

1: ALSPAC

Male

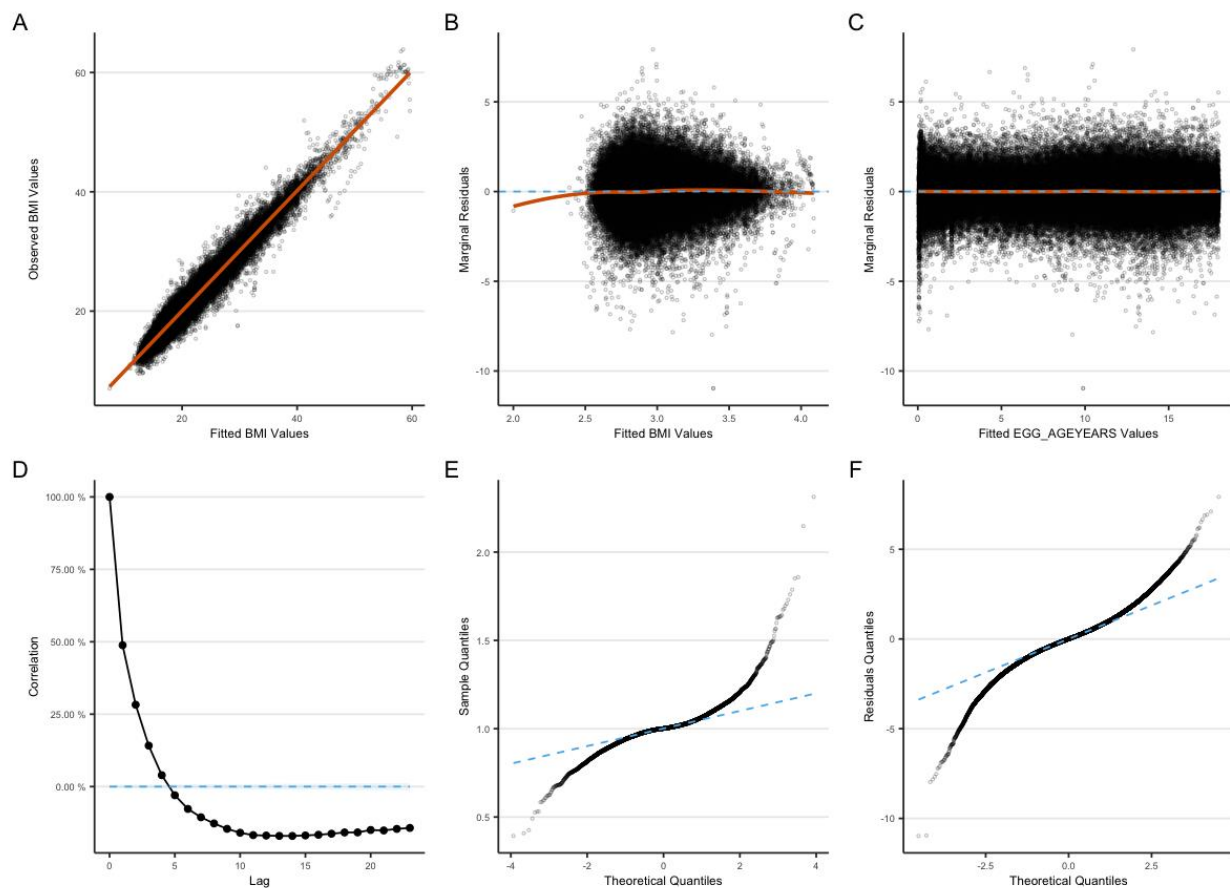


Female



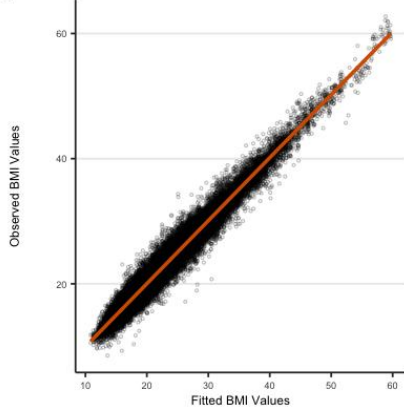
2: CHOP-EA

Male

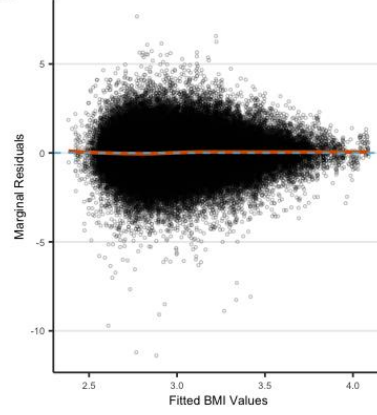


Female

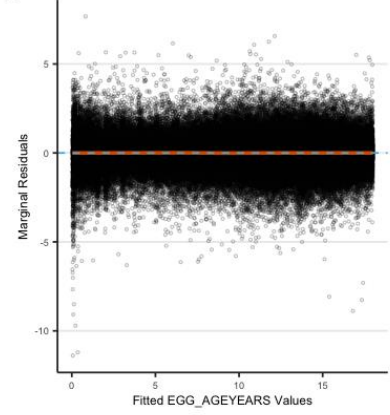
A



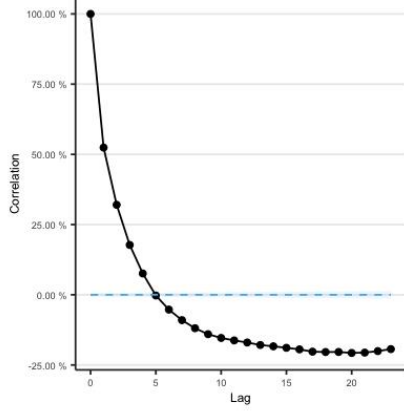
B



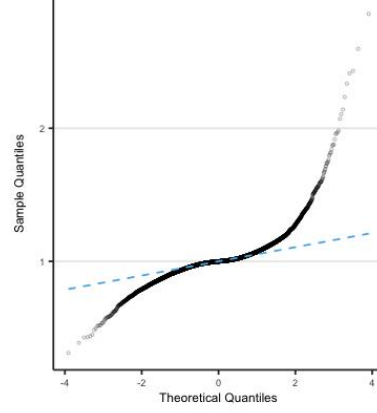
C



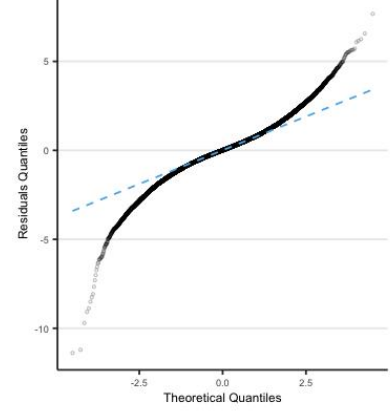
D



E

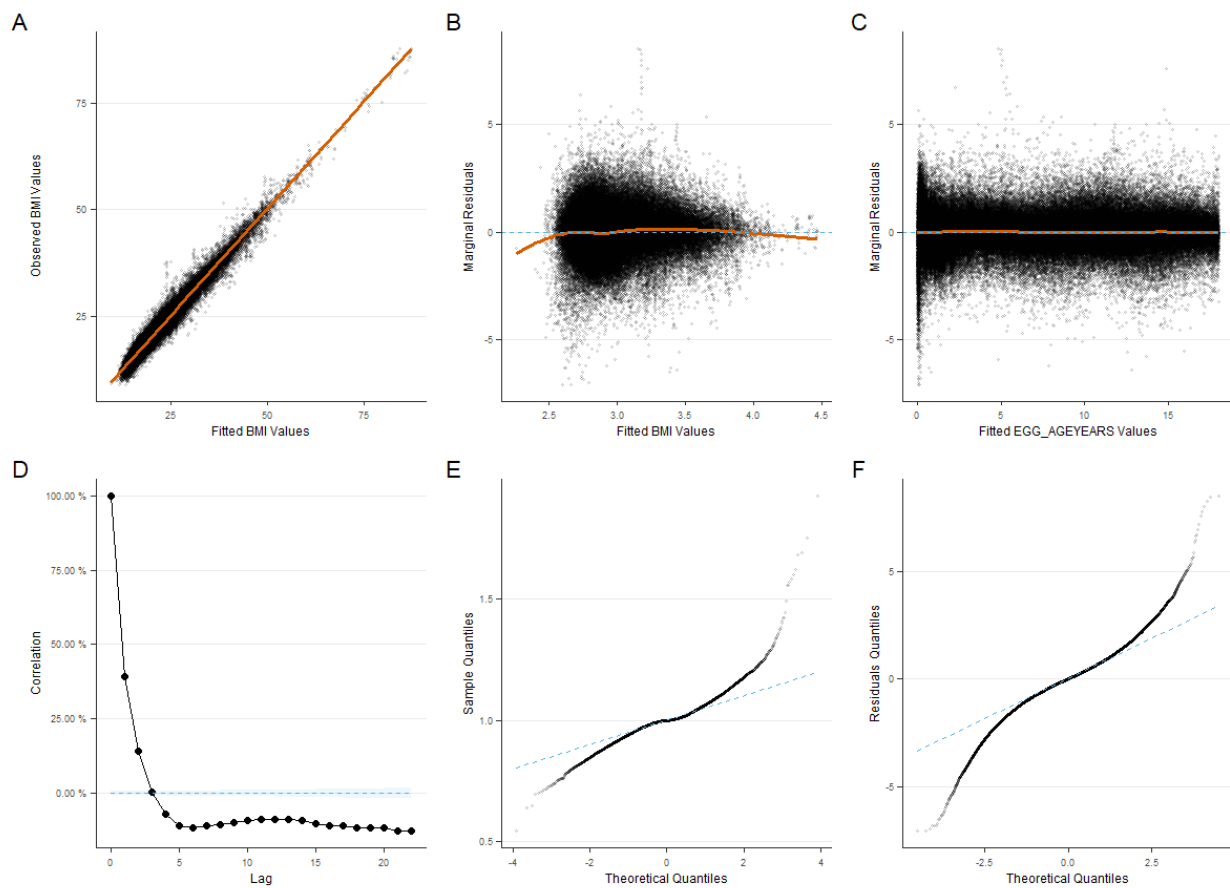


F



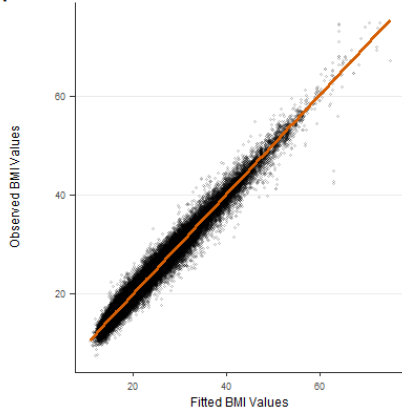
3: CHOP-AA

Male

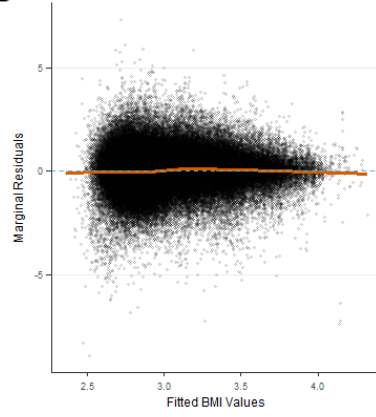


Female

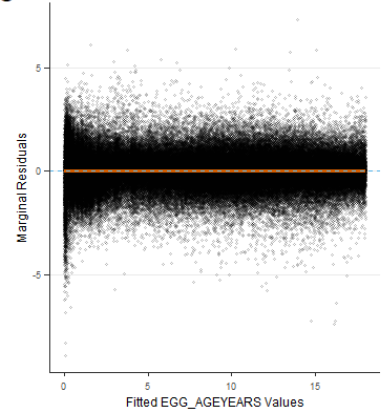
A



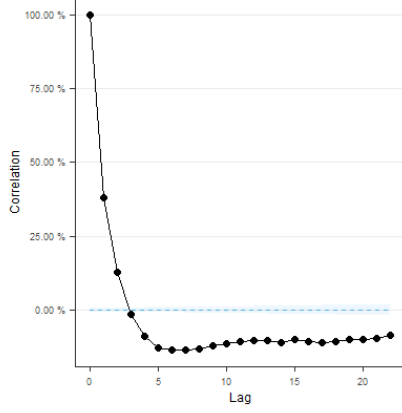
B



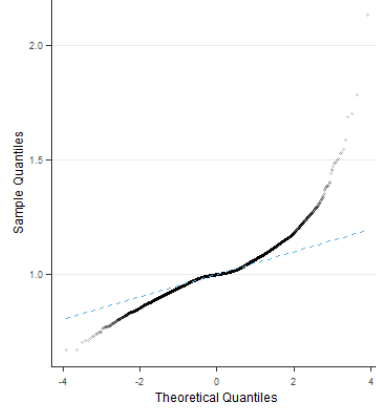
C



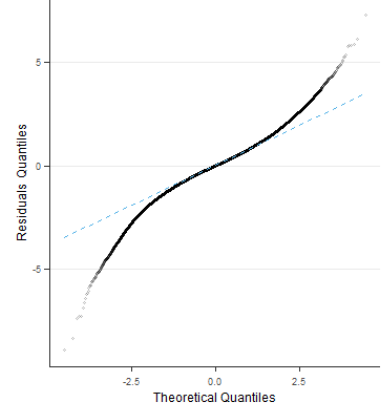
D



E

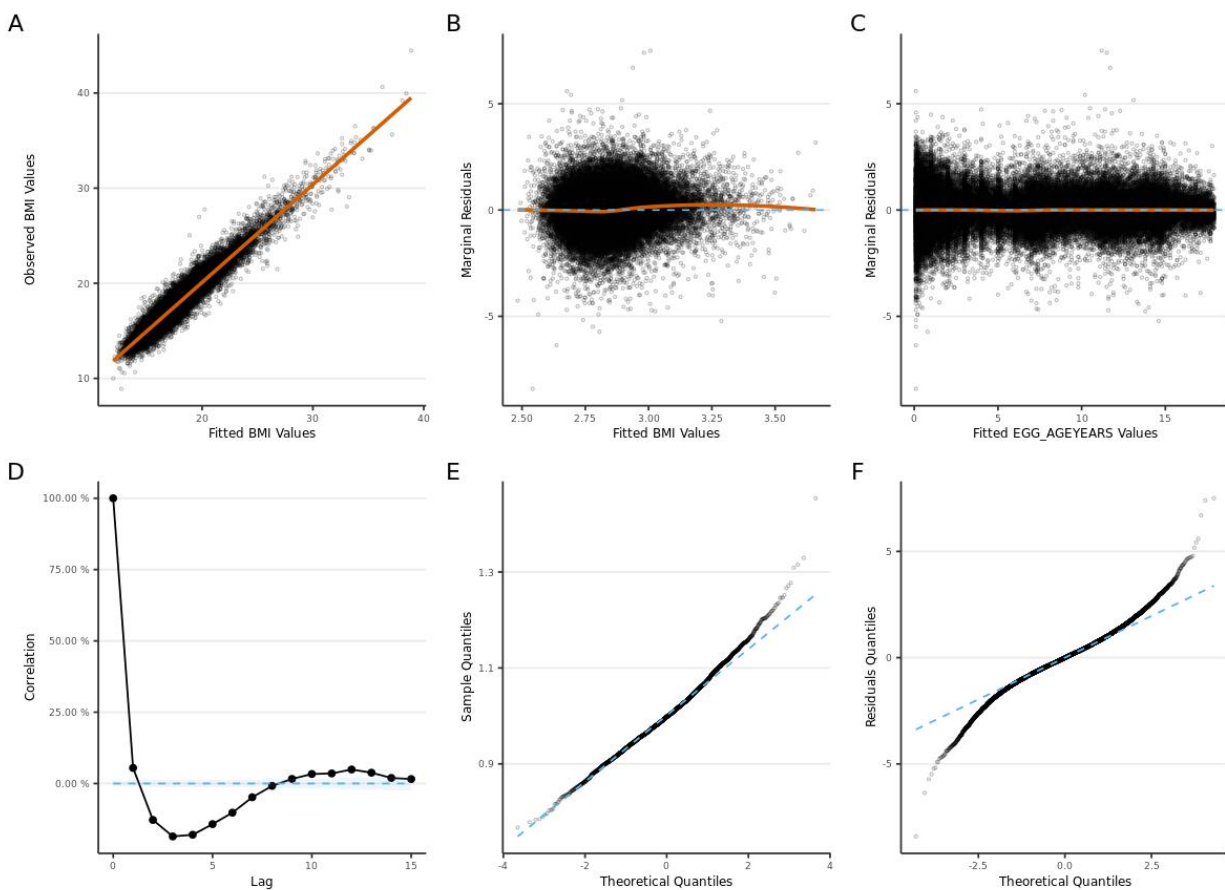


F



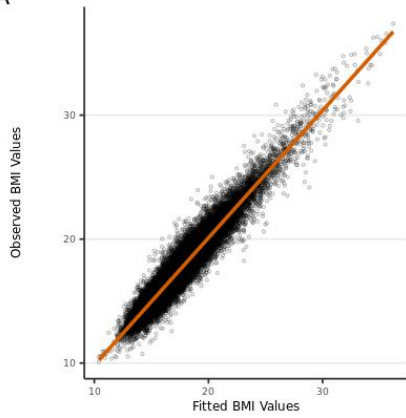
4: NFBC1966

Male

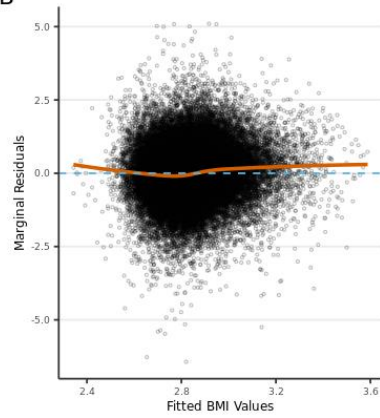


Female

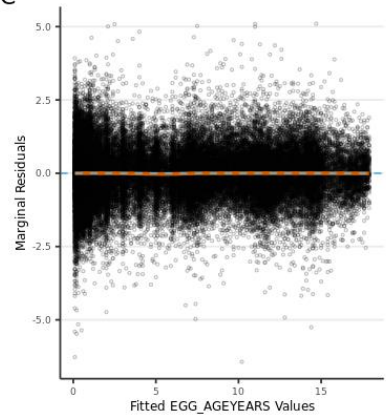
A



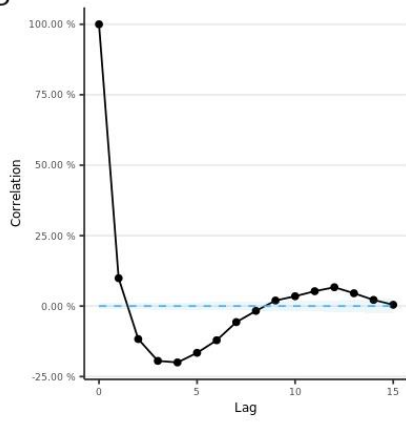
B



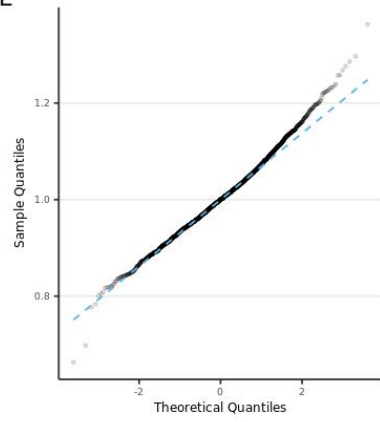
C



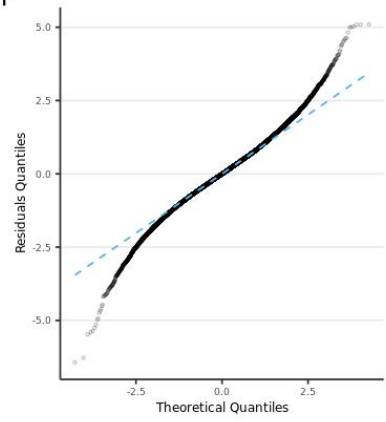
D



E

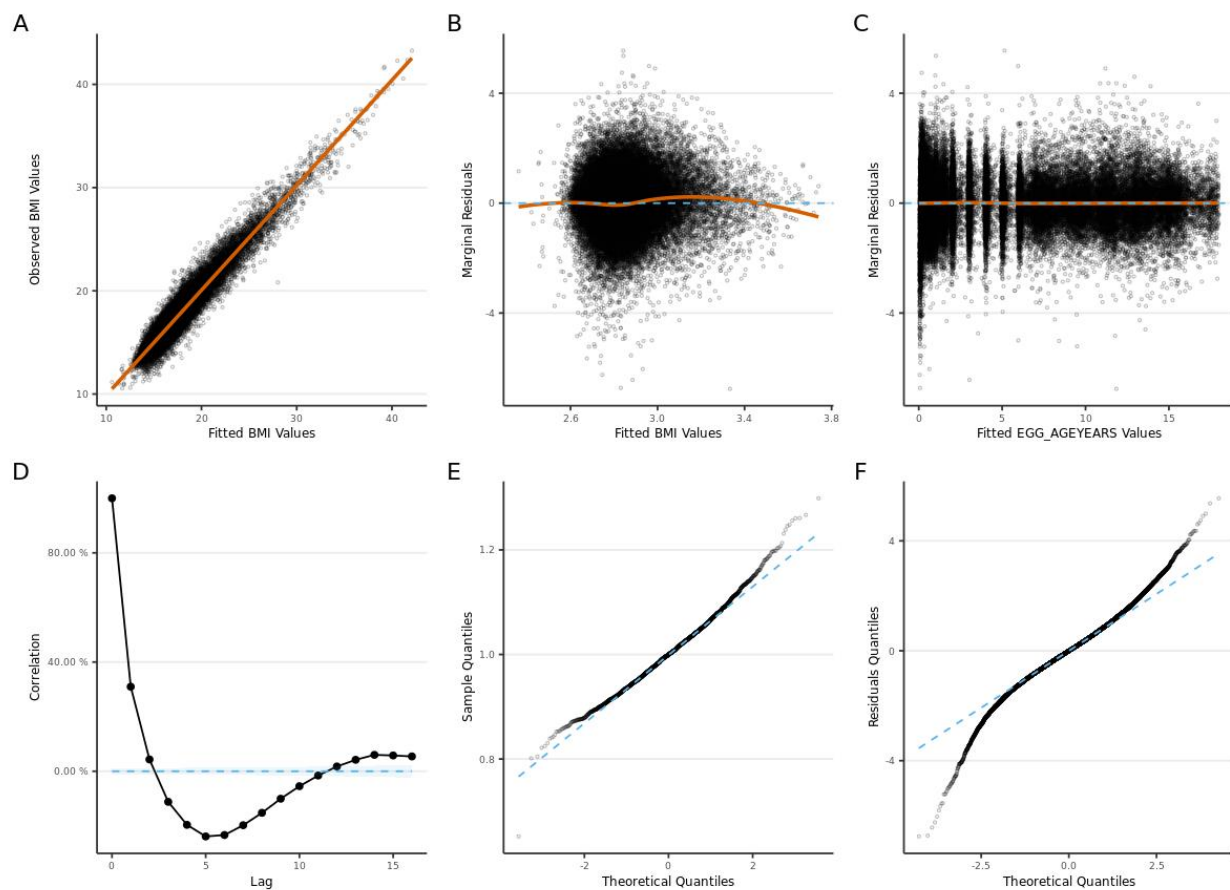


F

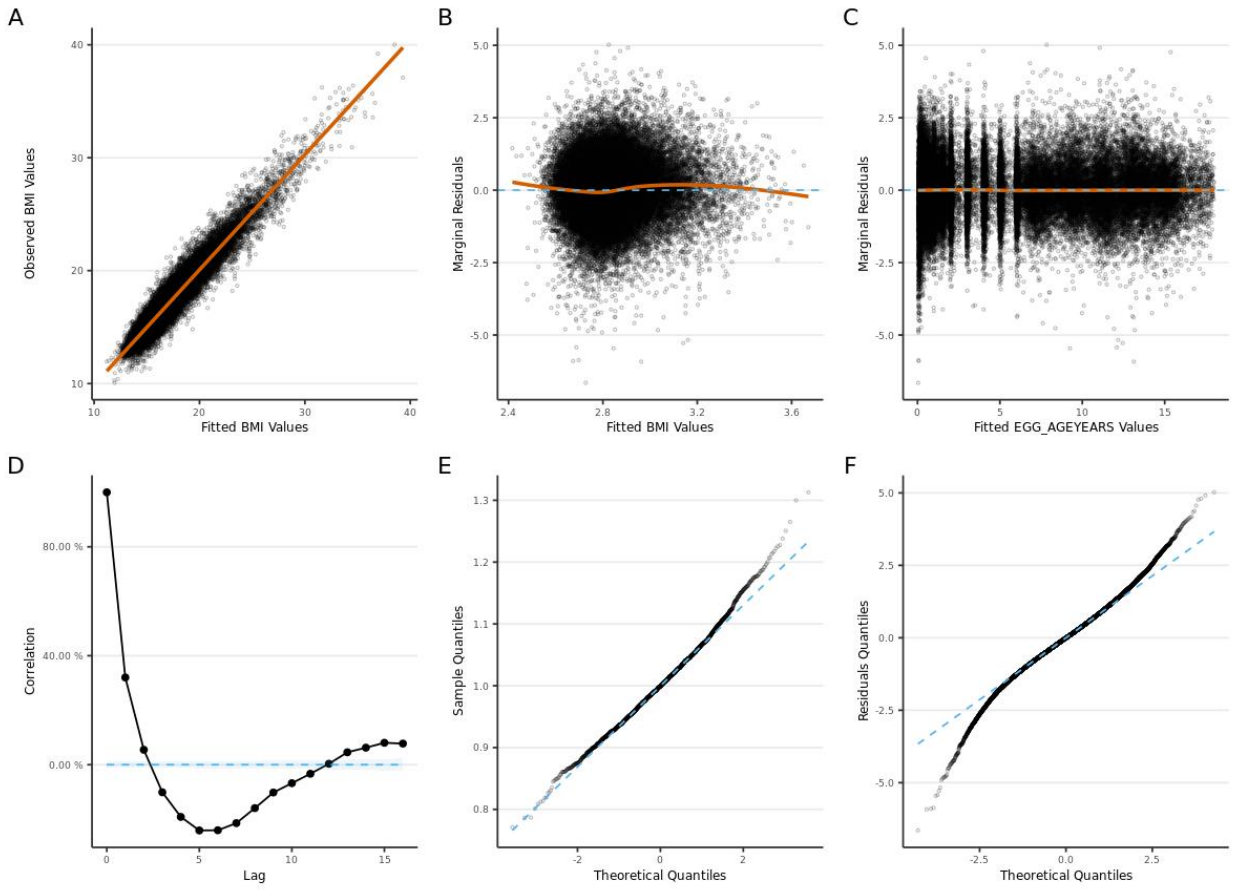


5: NFBC1986

Male

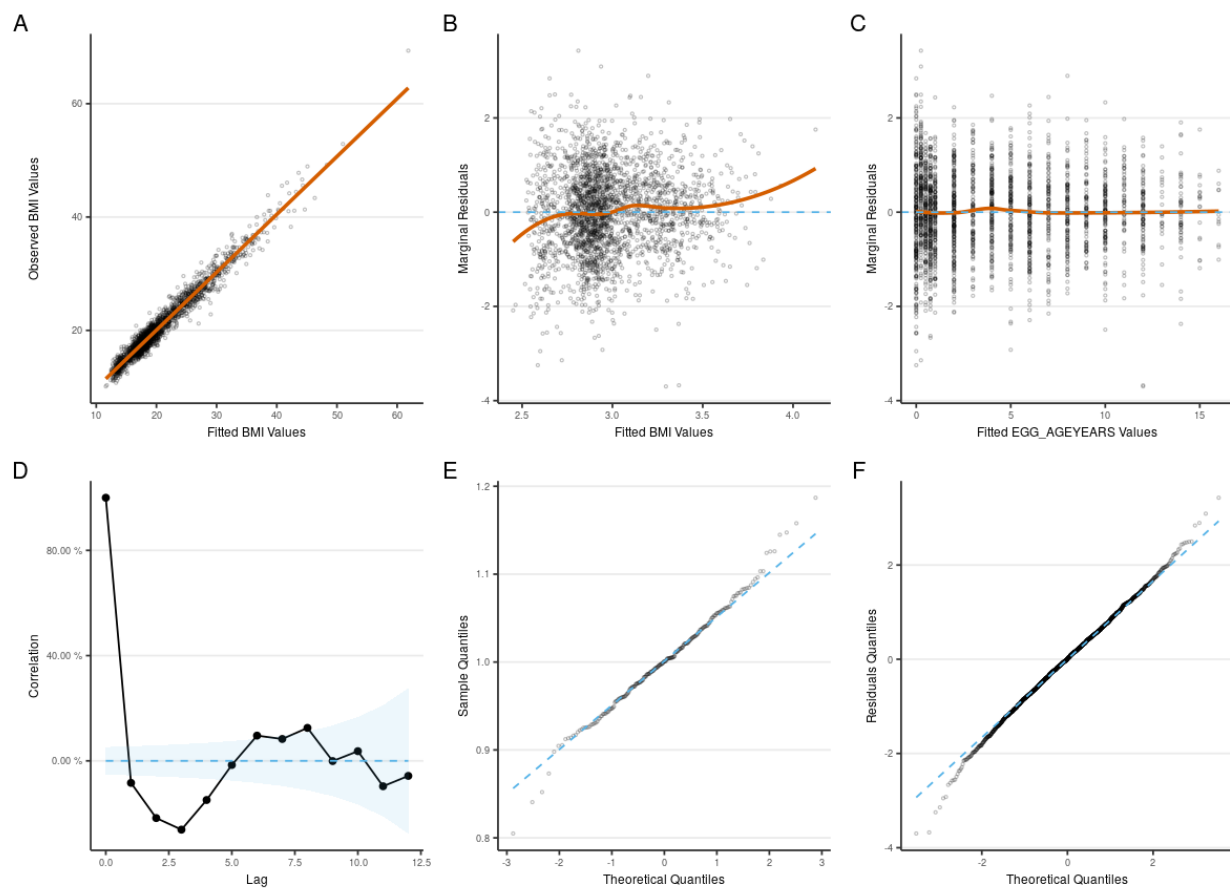


Female

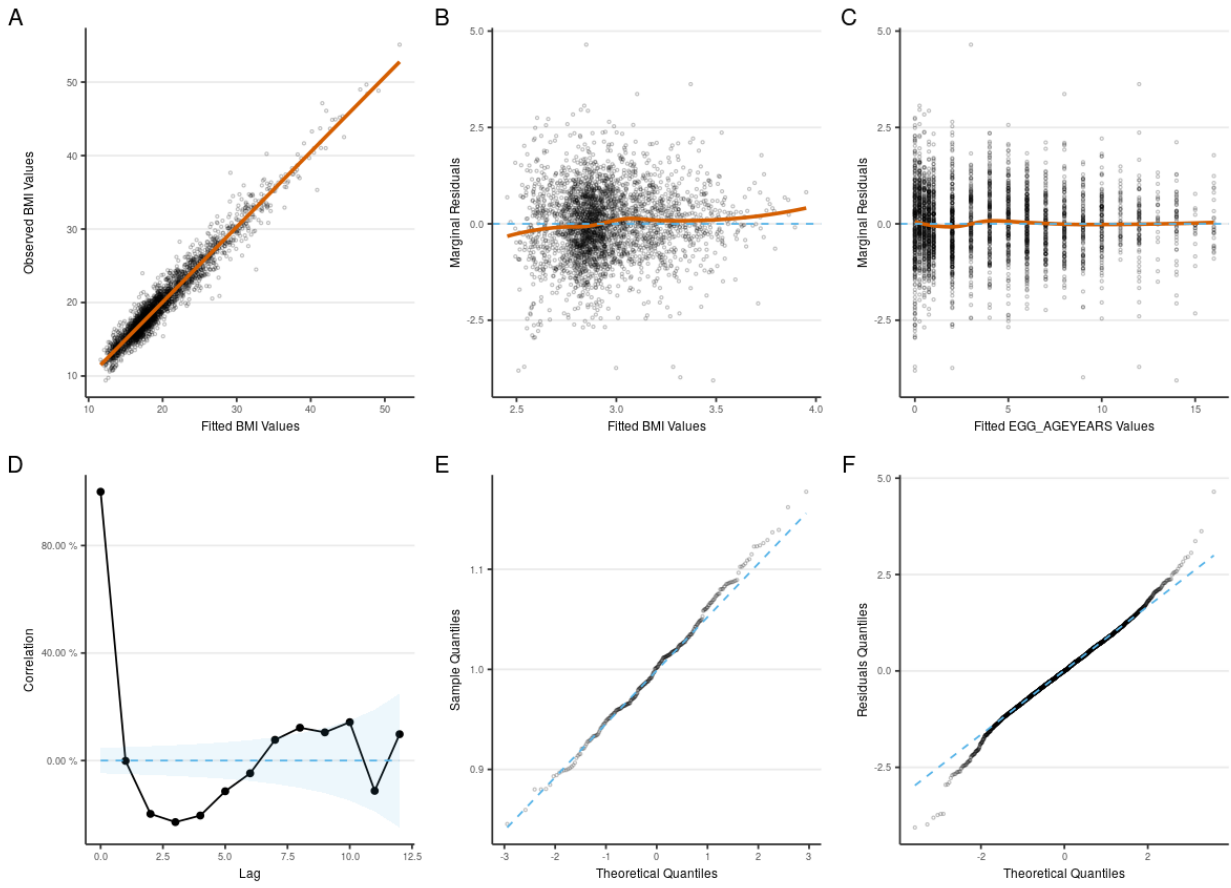


6: OBE

Male

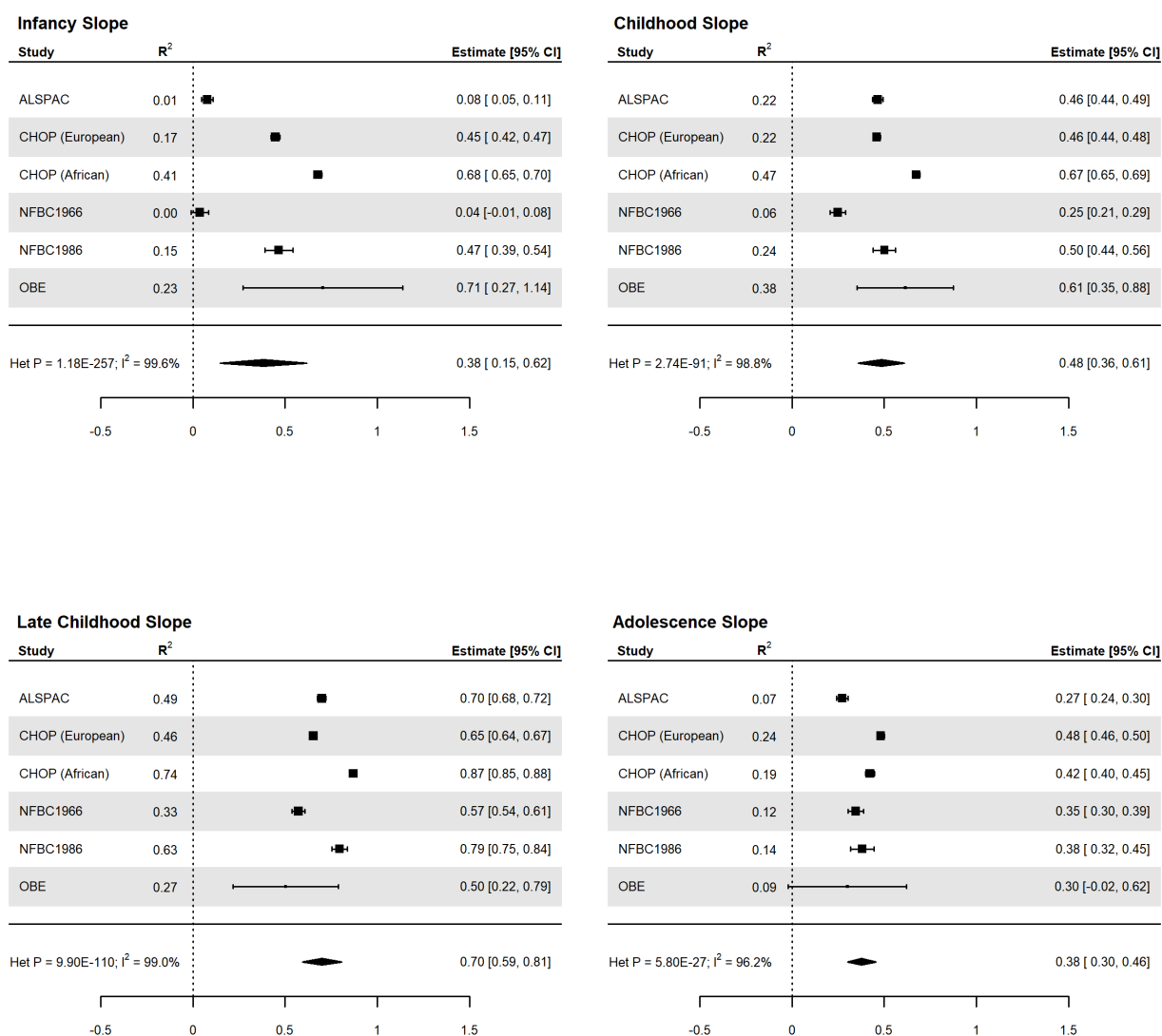


Female

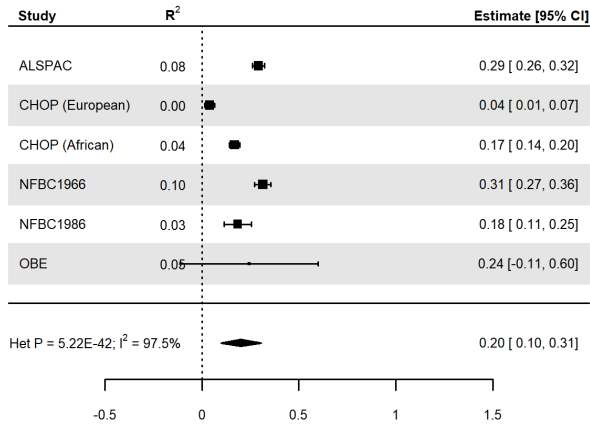


Supplementary Figure 4: Summary of the association between final BMI (between age 16-18 years) and each of the estimated phenotypes, including a random effects meta-analysis (DerSimonian-Laird estimator).

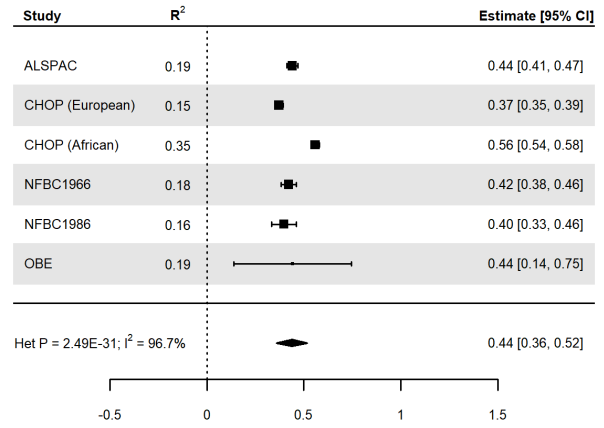
Within each cohort, both final BMI and the estimated phenotypes were standardized and analyses were adjusted for sex.



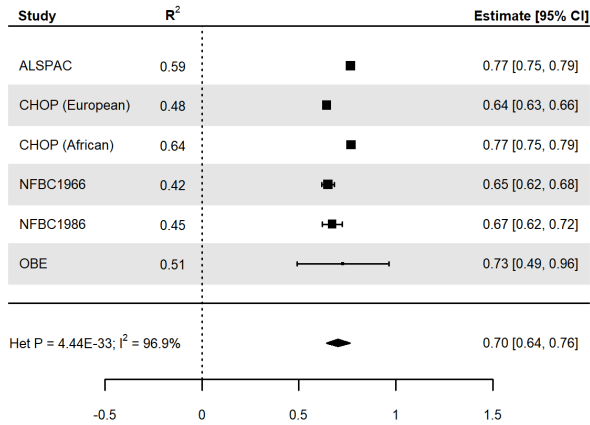
Infancy AUC



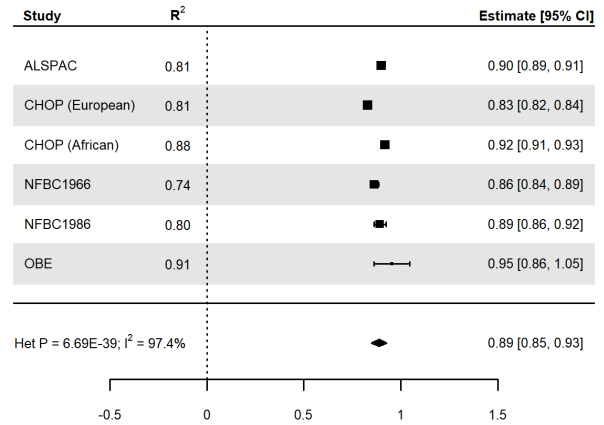
Childhood AUC



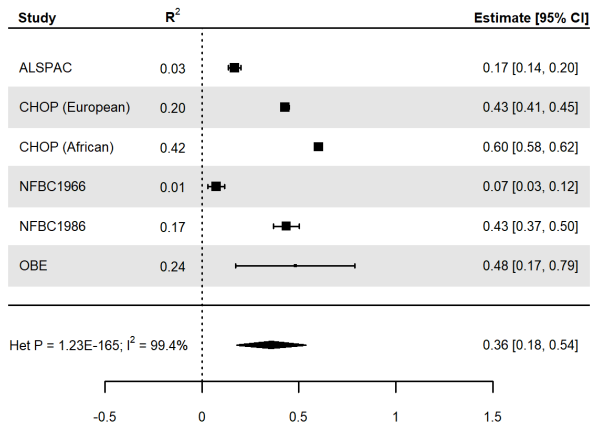
Late Childhood AUC



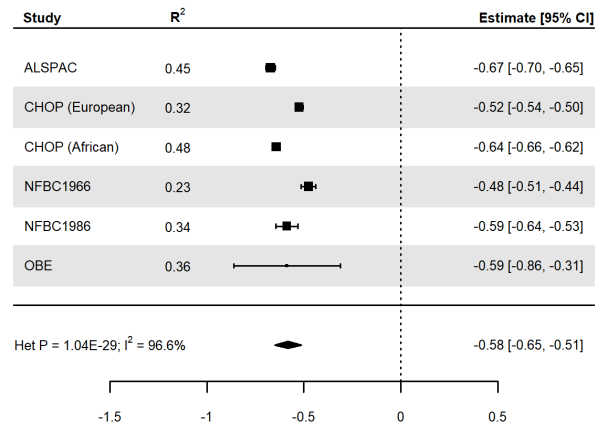
Adolescence AUC



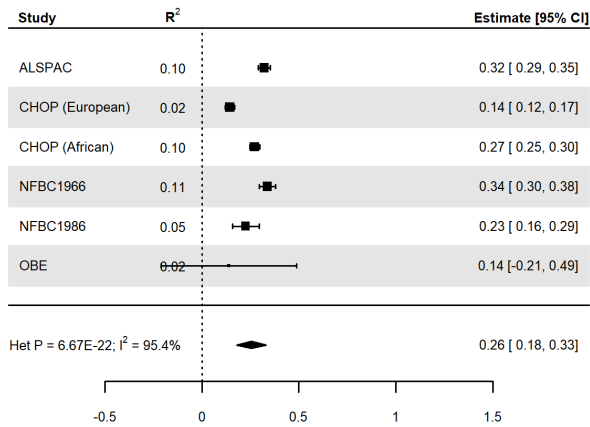
Age at AP



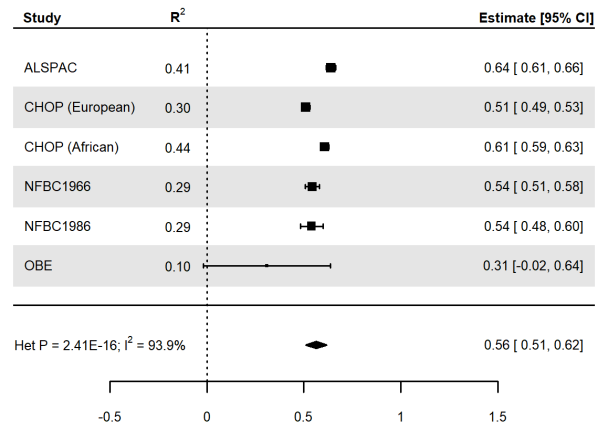
Age at AR



BMI at AP

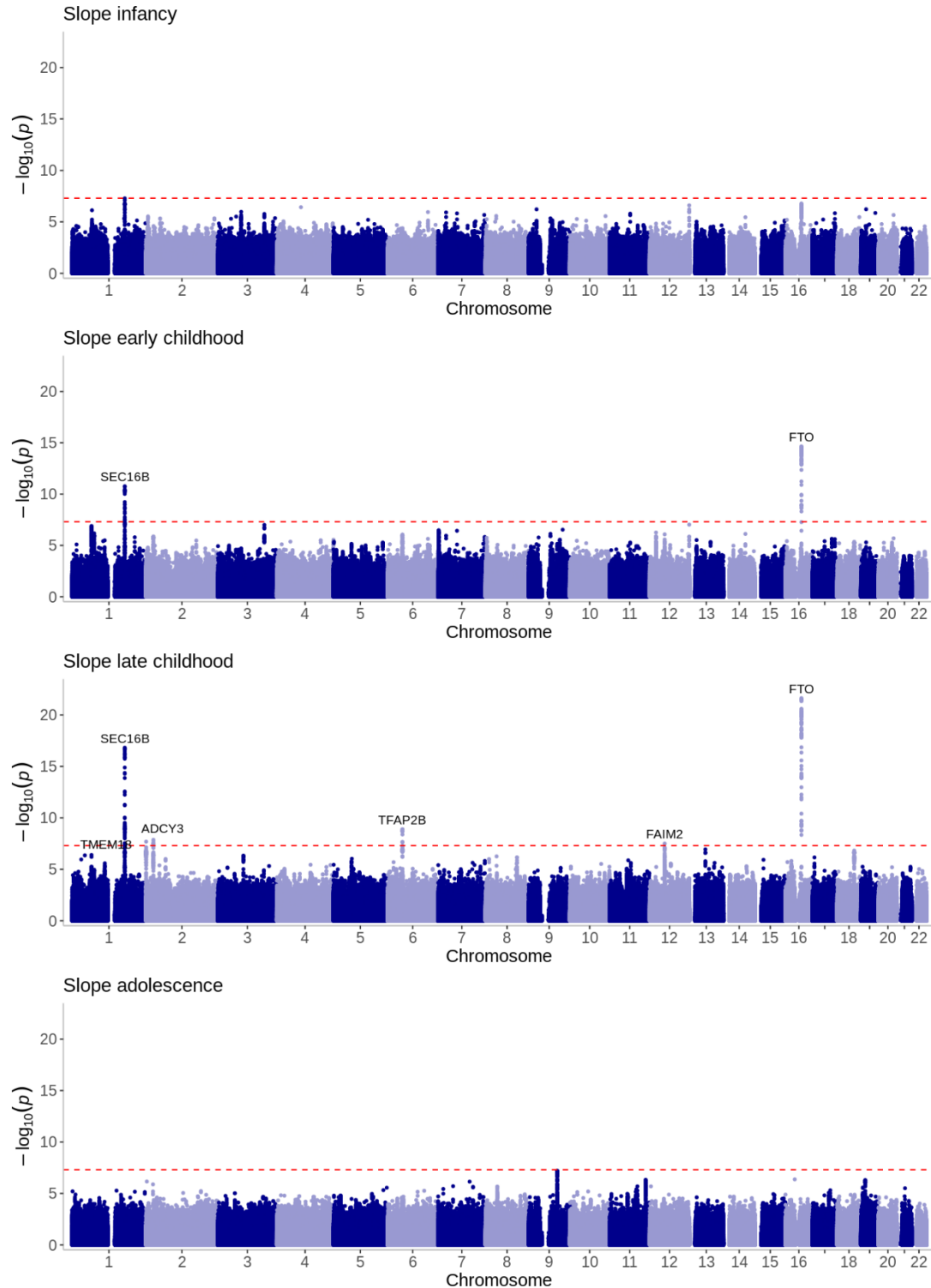


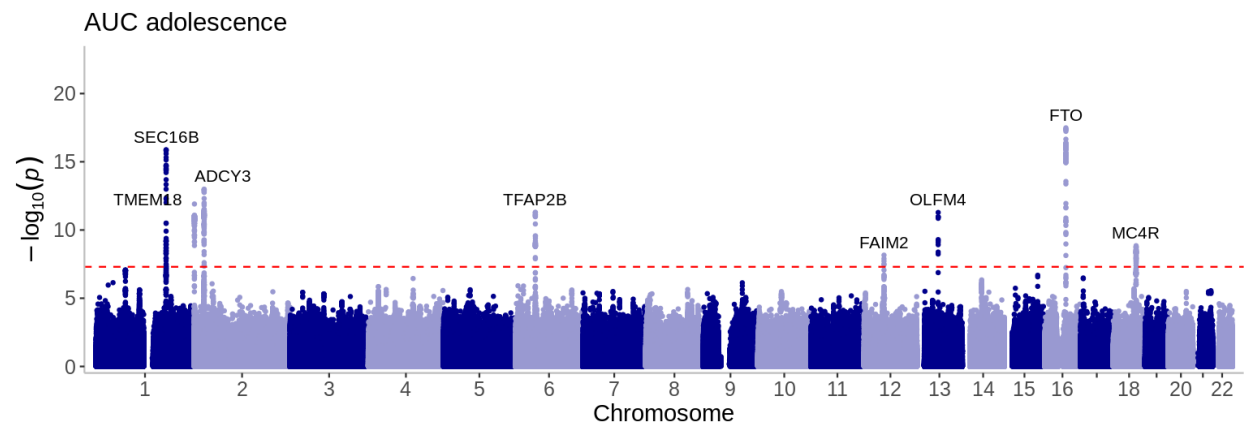
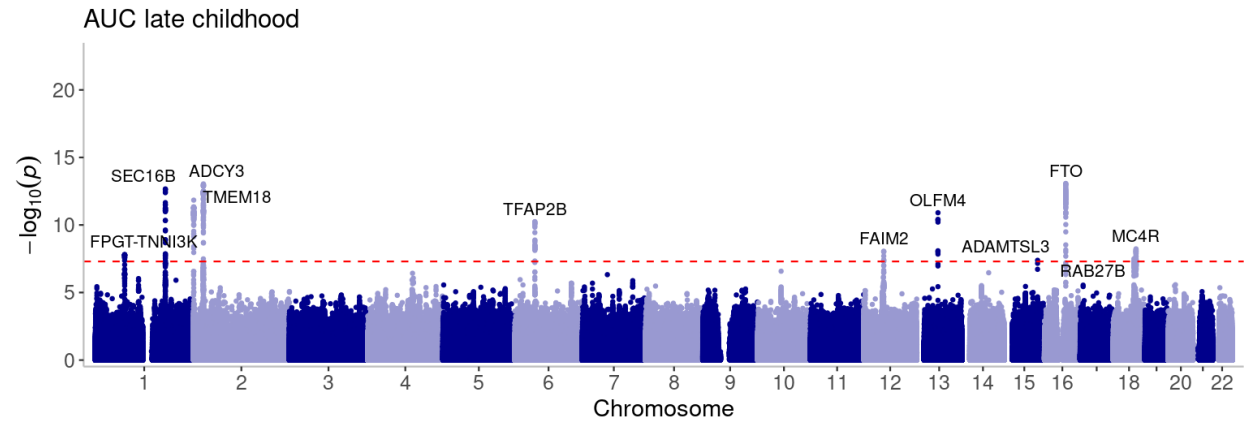
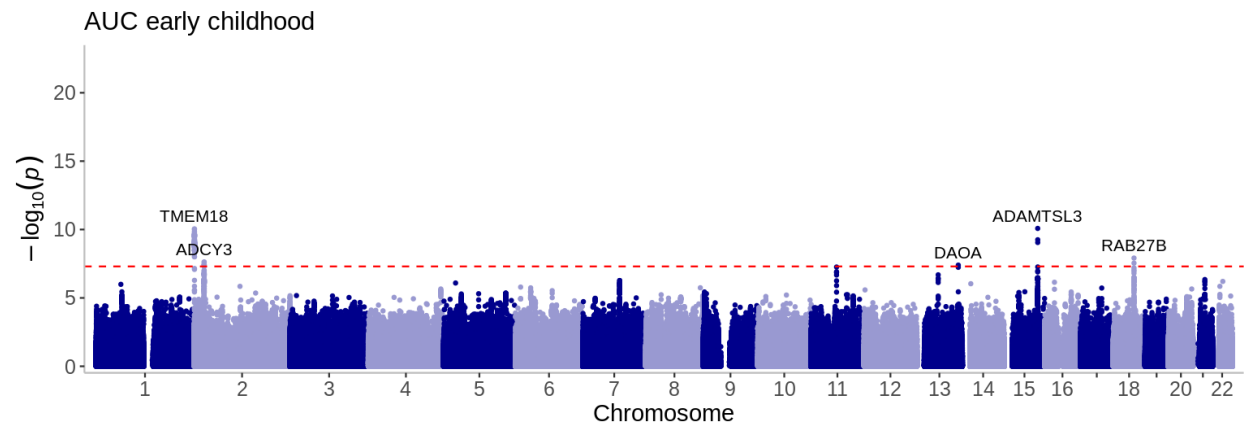
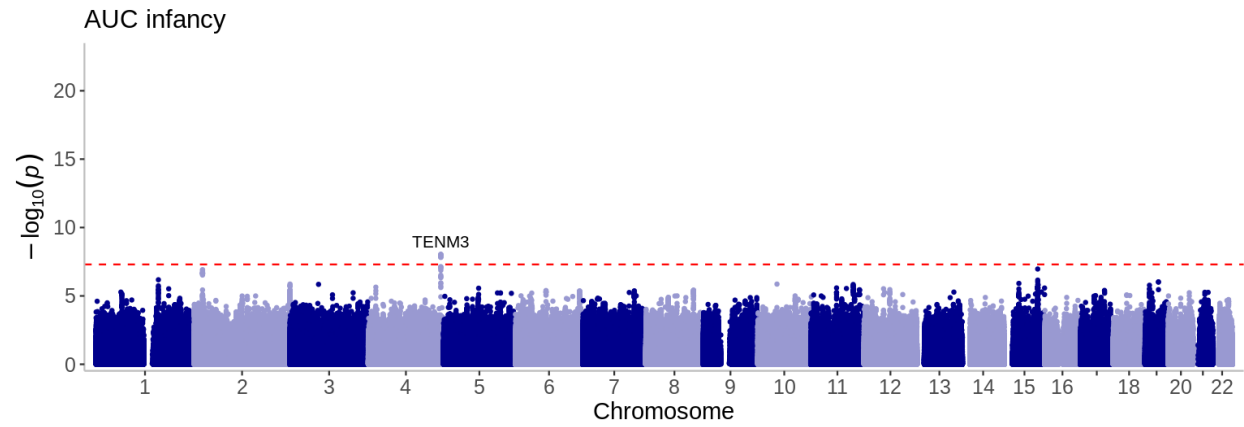
BMI at AR

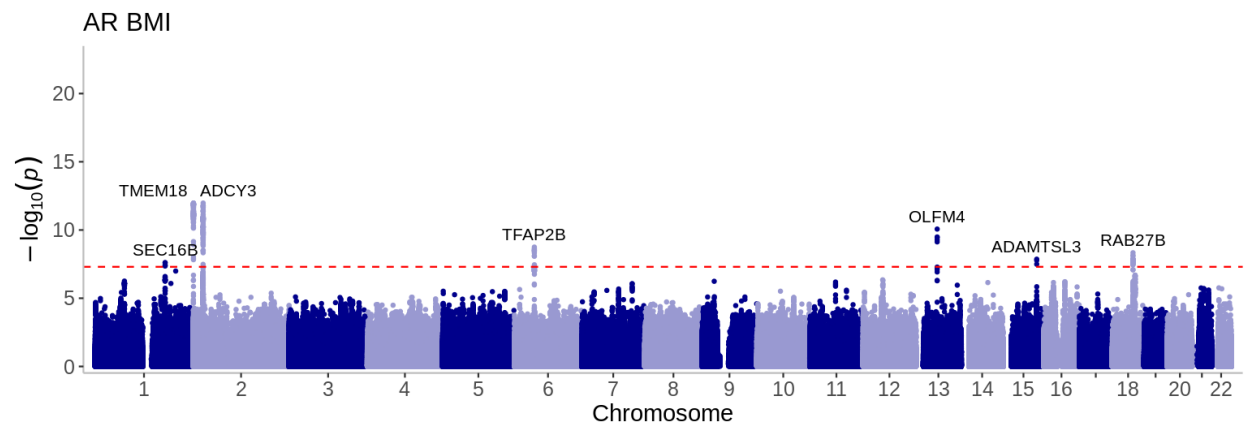
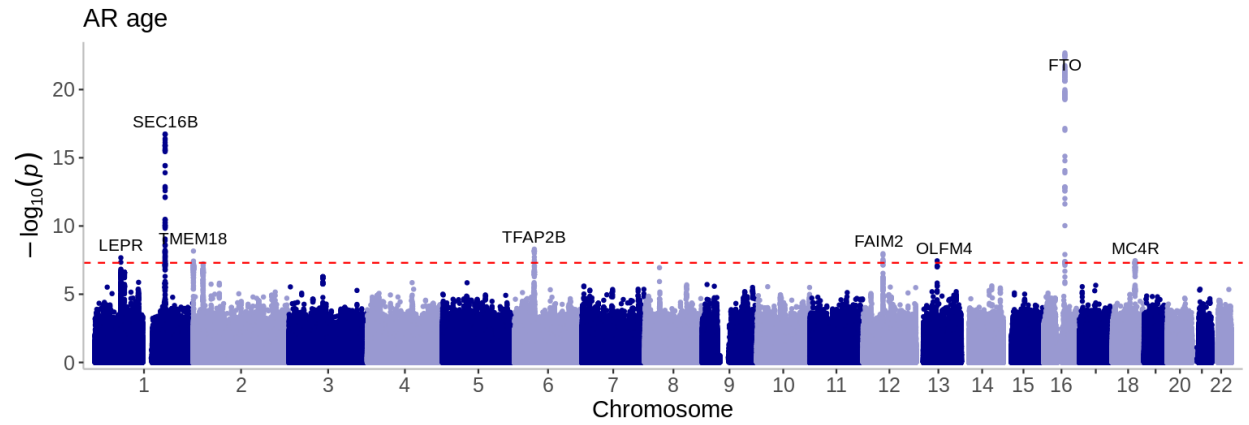
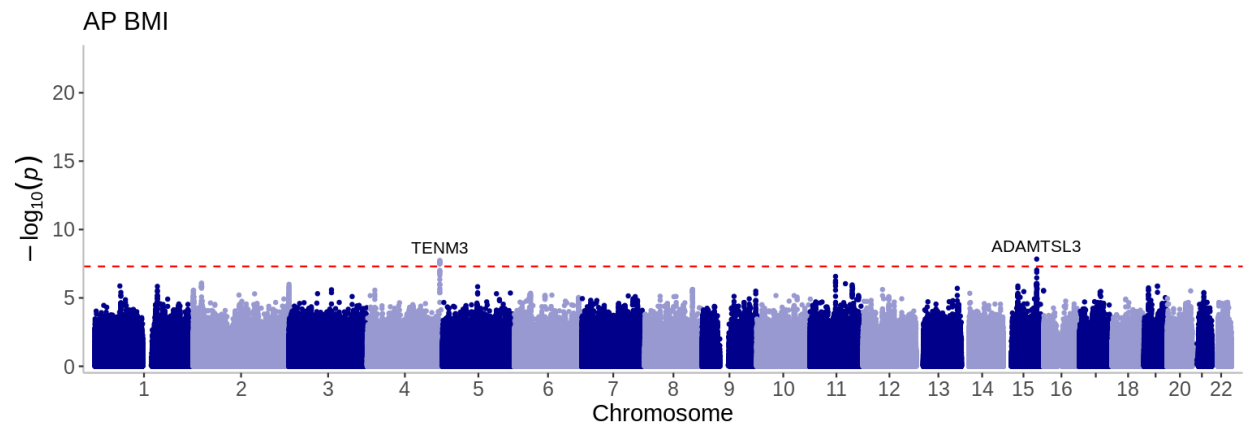
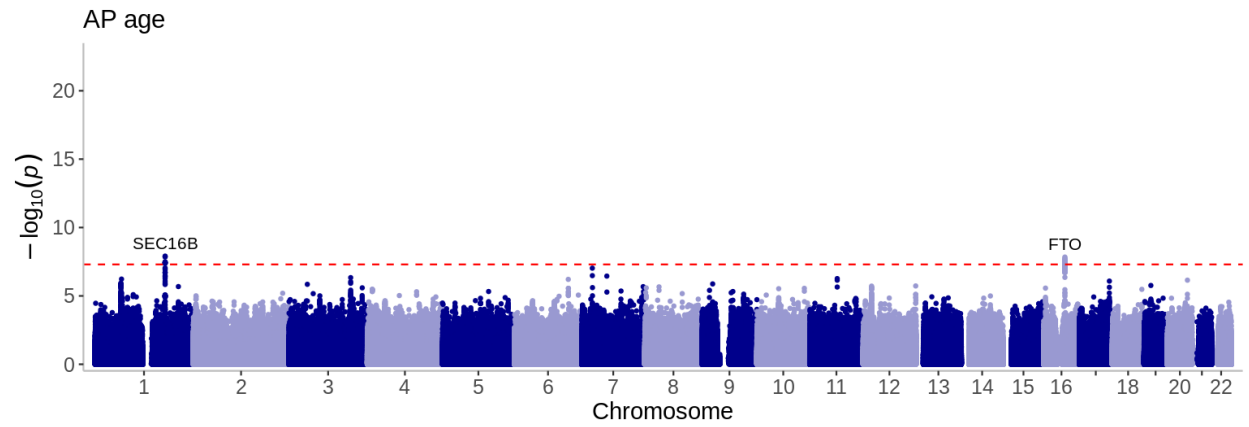


Supplementary Figure 5: Manhattan plots of the meta-analyses without OBE for the phenotypes across infancy (0-0.5 years), early childhood (1.5-3.5 years), late childhood (6.5-10 years) and adolescence (12-17 years).

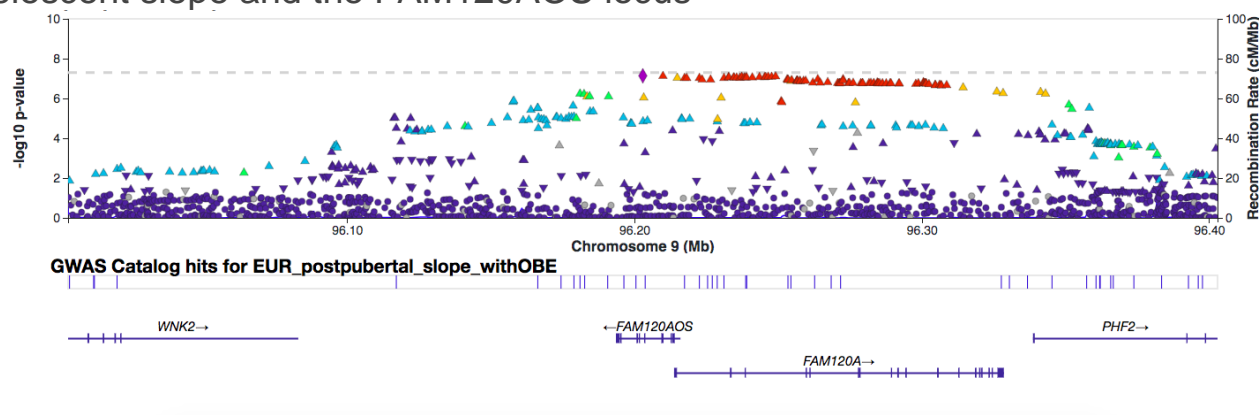
Loci are labelled with their nearest gene annotated by LocusZoom. The red line corresponds to the genome-wide significance level of $P < 5 \times 10^{-8}$







Supplementary Figure 6: LocusZoom regional plot of the association between adolescent slope and the FAM120AOS locus



Cohort description, acknowledgements and funding

Avon Longitudinal Study of Parents and Children (ALSPAC)

Pregnant women resident in Avon, UK with expected dates of delivery between 1st April 1991 and 31st December 1992 were invited to take part in the study (Boyd et al. 2013; Fraser et al. 2013). The initial number of pregnancies enrolled was 14,541. Of the initial pregnancies, there was a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age. When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above: The number of new pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented in the released data and reflecting enrolment status at the age of 24 is 906, resulting in an additional 913 children being enrolled (456, 262 and 195 recruited during Phases II, III and IV respectively). The total sample size for analyses using any data collected after the age of seven is therefore 15,447 pregnancies, resulting in 15,658 fetuses. Of these 14,901 children were alive at 1 year of age.

Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool:

<http://www.bristol.ac.uk/alspac/researchers/our-data/>

From birth to approximately five years of age, length and weight measurements were abstracted from health visitor records. Health visits form part of the standard childcare in the UK and data was abstracted from records at approximately six weeks, 10, 21, and 48 months of age. From ages 4 months to 5 years a sub-study (called “Child in Focus”) included a random 10% of the cohort and recorded data across eight research clinics. Length/height and weight were also measured at these clinics. From age 7 onwards, all children were invited to annual research clinics in which height and weight were measured. Furthermore, parent-reported height and weight was also recorded via questionnaires across the time course. BMI was derived from the length/height and weight measures (mean of 9 measures per participant) as weight (kg) divided by the square of height (m). Given the inclusion of measured and self-reported height and weight data, we additionally included an indicator of the source of data as a covariable in the LMMs (clinic/health visit vs. questionnaire). All participants with at least one measure of BMI was included in the study, however, we excluded those who were part of a multiple birth (i.e. twins, triplets etc. n=524).

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). Genome-wide genotyping data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. This publication is the work of the authors and Kimberley Burrows will serve as guarantors for the contents of this paper.

For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Northern Finland Birth Cohorts born in 1966 (NFBC1966)

The Northern Finland Birth Cohort study 1966 (NFBC1966) includes 12,058 live born individuals, of European descent, with expected dates of birth during 1966 in the two northernmost provinces of Finland, Oulu and Lapland (1966 University of Oulu. 1966). The data on all cohort members were prospectively collected since pregnancy and supplemented at the ages of 1, 14, 31 and 46 years (Nordström et al. 2021). Height and weight growth measurements were obtained from communal child health clinics (Sovio et al. 2009). All those living in northern Finland or in the capital area were invited to a clinical examination and blood sampling at age 31 years. DNA was extracted from 5753 individuals and Illumina's HumanCNV370-Duo DNA Analysis BeadChip was used to obtain genome-wide genetic data (Sabatti et al. 2009).

Prior to growth trajectory modelling, we excluded participants born pre-term (n=765) and non-singletons (n=197). We also identified the sibling pairs of this cohort that were born at different times, and one participant per sibling pair was excluded (n=17).

Approval for the studies was granted by the Northern Ostrobothnia Hospital District Ethical Committee 94/2011 (12.12.2011), Finland in accordance with the declaration of Helsinki. Mothers gave their informed consent in the beginning of the NFBC1966 data collection. Written informed consent has been obtained from the cohort participants in the 31- and 46-year data collections.

We thank all cohort members and researchers who participated in the study. We also wish to acknowledge the work of the NFBC project center. NFBC1966 31-year follow-up received financial support from University of Oulu Grant no. 65354, Oulu University Hospital Grant no. 2/97, 8/97, Ministry of Health and Social Affairs Grant no. 23/251/97, 160/97, 190/97, National Institute for Health and Welfare, Helsinki Grant no. 54121, Regional Institute of Occupational Health, Oulu, Finland Grant no. 50621, 54231. NFBC1966 46yr follow-up received financial support from University of Oulu Grant no. 24000692, Oulu University Hospital Grant no. 24301140, ERDF European Regional Development Fund Grant no. 539/2010 A31592.

Northern Finland Birth Cohorts born in 1986 (NFBC1986)

The Northern Finland Birth Cohort 1986 (NFBC1986) includes 9,432 live born children with expected dates of birth between 1st July 1985 and 30th June 1986 in the two northernmost provinces of Finland, Oulu and Lapland (Järvelin, Hartikainen-Sorri, and Rantakallio 1993; Järvelin et al. 1997). The cohort has been followed up since early pregnancy until adulthood (1986 University of Oulu. 1986). Growth measurements were obtained from communal child health clinics. All those alive with known address were invited to a clinical examination at the age of 15 to 16 years. At this age, blood samples were drawn and DNA was extracted for 6,266 subjects.

Prior to growth trajectory modelling, we excluded pre-term babies (n=496) as well as twins and triplets (n=132). From the remaining sample, we further identified one sibling pair that was born at different times, and the child with less height and weight measurements was excluded. Finally, one child with extremely fast BMI growth from early on was excluded from the sample as the trajectory was not comparable to other children.

Approval for the studies was granted by the Northern Ostrobothnia Hospital District Ethical Committee 108/2017 (15.1.2018), Finland in accordance with the declaration of Helsinki. Participants provided written informed consent in data collections carried out when they were aged 15 to 16 and 33 to 35 years.

We thank all cohort members and researchers who have participated in the study. We also wish to acknowledge the work of the NFBC project center. The cohort received financial support from the following grants EU QLGI-CT-2000-01643 (EUROBLCS) Grant no. E51560, NorFA Grant no. 731, 20056, 30167, USA / NIH 2000 G DF682 Grant no. 50945.

CHOP

The CHOP cohort is a random sampling of the population of children who come for care at the Children's Hospital of Philadelphia starting in 2006 and continuing presently (Connolly et al. 2020). If patients choose to enroll in the study, their fully de identified and encrypted electronic medical records are available for study. Heights and weights used to calculate BMI in this study were obtained from electronic medical records. Self-reported ancestry was used to define African Americans and Europeans Americans. The CHOP study was genotyped at the Center for Applied Genomics at the Children's Hospital of Philadelphia. The Research Ethics Board of CHOP approved the study, and written informed consent was obtained from all subjects.

The authors thank the network of primary care clinicians and the patients and families for their contribution to this project and to clinical research facilitated by the Pediatric Research Consortium [PeRC]-The Children's Hospital of Philadelphia. R. Chiavacci, E. Dabaghyan, A. [Hope] Thomas, K. Harden, A. Hill, C. Johnson-Honesty, C. Drummond, S. Harrison, F. Salley, C. Gibbons, K. Lilliston, C. Kim, E. Frackelton, F. Mentch, G. Otieno, K. Thomas, C. Hou, K. Thomas and M.L. Garriss provided expert assistance with genotyping and/or data collection and management. The authors would also like to thank S. Kristinsson, L.A. Hermannsson and A. Krisbjörnsson of Raförinn ehf for extensive software design and contributions. This research was financially supported by an Institute Development Award from the Children's Hospital of Philadelphia, a Research Development Award from the Cotswold Foundation, the Daniel B. Burke Endowed Chair for Diabetes Research, the Children's Hospital of Philadelphia Endowed Chair in Genomic Research and NIH grant R01 HD056465.

OBE

The OBésité de l'Enfant (OBE) cohort includes children with early-onset, familial obesity, who were born in France. Most of them were recruited by the CNRS 8199 unit in 1998. A small part of

these children were patients of Toulouse Children's Hospital. Height and weight measurements were taken by the participant's GP or pediatrician. The study protocols were approved by local ethics committees. Oral assent from children was obtained and parents (or legal guardians) signed an informed consent form.

The authors would like to thank the French National Research Agency (Agence Nationale de la Recherche [ANR]-10-LABX-46 [European Genomics Institute for Diabetes] and ANR-10-EQPX-07-01 [Lille Integrated Genomics Advanced Network for personalized medicine] to A.B. and P.F.), European Research Council (OpiO 101043671 to A.B.), the European Union's Horizon Europe Research and Innovation Programme (OBELISK grant agreement 101080465 to A.B. and P.F.), and the National Center for Precision Diabetic Medicine (PreciDIAB to A.B. and P.F.), which is jointly supported by the French National Agency for Research (ANR-18-IBHU-0001), European Regional Development Fund, Hauts-de-France Regional Council, and the European Metropolis of Lille. The authors also thank the France Génomique consortium (ANR-10-INBS-009).

Linear mixed modelling

A simple approach is to model a repeated continuous outcome (e.g. BMI) as a function of time within a linear mixed-effects model (LMM). LMM has an advantage over other methods, such as repeated measures ANOVA, by allowing missing data within individuals, and unbalanced study designs where either the number of measures or the timing of the measurements differs between individuals, as well as fitting and testing covariance structures (Laird and Ware 1982).

A basic random-slope LMM with a single continuous outcome such as BMI as a linear function of time can be written as:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + v_{0i} + v_{1i} X_{it} + \varepsilon_{it}$$

Where Y_{it} is a single outcome measured for individual i at time t and is assumed to be independent between individuals. The fixed coefficients, β_0 and β_1 represent the average intercept and slope respectively. The random coefficients represent the deviation from the average intercept (v_{0i}) and slope (v_{1i}) for individual i . The random effects are assumed to follow a bivariate normal distribution with mean zero and covariance matrix Ω_v . Residual errors ε_{it} are assumed to be independently identically normally (i.i.d) distributed with variance $\sigma_{\varepsilon 0}^2$.

The LMM model above assumes a linear change in the continuous outcome with increasing time. However, growth is nonlinear for many biological characteristics, such as BMI across early life. BMI throughout infancy, childhood, and adolescence follows a complex pattern of change including the appearance of the adiposity peak during infancy and then a rebound during childhood (see Figure 1 in the main text). This nonlinear change can be incorporated into LMM models through the addition polynomial functions of time. A quadratic function for instance can induce a curve with a single turn, while a cubic function can involve up to two turns. More simple polynomials may not adequately describe the turning points in the underlying data, while more complex polynomials can produce artefactual turns in the curve and can particularly fit poorly at extremes where data is more sparse. A more flexible approach to model complex curves such as BMI is through using spline functions, joined at knot points dispersed across time. Smooth functional polynomials (usually of low order) are chosen to fit the measures between consecutive knots (Perperoglou et al. 2019).

Example process and output from the *EGGLA* BMI modelling framework

While the framework described below is specific to the current study exploring the GWAS of longitudinal BMI trajectories and estimated phenotypes, the framework can serve as an exemplar for other consortium efforts to analyse other non-linear longitudinal traits.

To undertake LMM analysis of BMI, analysts are required to load in longitudinal data in long format with the following mandatory columns: individual ID, sex, height measurement (cm), weight measurement (kg), and age at measurement (days or years). Additional cohort specific covariates may also be included.

A comprehensive protocol has been written to allow the user to set up the environment and work through the analysis pipeline. This is available at: <https://m.canouil.dev/eg gla/articles/eg gla.html>. The online protocol contains a series of vignettes containing the code used to perform the model diagnostics framework (<https://m.canouil.dev/eg gla/articles/models-diagnostics.html>), a plot to aid in model selection (<https://m.canouil.dev/eg gla/articles/model-selection.html>), analyses and plots specific to the chosen model from the diagnostics framework (in this case specific to the cubic spline model) (<https://m.canouil.dev/eg gla/articles/run-cubic-splines.html>), and finally characterisation of the AP and AR (<https://m.canouil.dev/eg gla/articles/adiposity-peak-rebound.html>). See each vignette for examples of output using simulated data.

The chosen best model selected from the output of the vignettes above was taken forward for model refinement (see methods of main paper). Finally, two *EGGLA* R package (Canouil et al. 2024) functions; `run_eg gla_lmm()` and `run_eg gla_gwas()` were used to perform the LMM and GWAS using the final chosen model (cubic spline function in the fixed effects with cubic slope function random effects).

The *EGGLA* (Canouil et al. 2024) function `run_eg gla_lmm()` is a wrapper used to perform the following analysis stages: 1) cleaning and formatting of BMI data using the *growthcleanr* R package for longitudinal data cleaning, 2) deriving BMI from cleaned height and weight data and formatting of the dataframe for analysis, 3) run the LMM for males and females separately, 4) plotting of model residuals, 5) deriving slopes and AUCs for defined linear time intervals of the curve, 6) predict BMI and derive age and BMI at AP and AR, 7) detection of outlier individuals for each of the estimated phenotypes and output of correlations between latent phenotypes, 8) formatting the derived phenotypes into a dataset ready for GWAS. The `run_eg gla_lmm()` function allows for further refinement of models where applicable including: addition of covariates, choice of knot points and placement, choice of the linear time periods for derivation of slopes and AUCs, inclusion of correlation structure and complexity of random effects. After specifying the parameters of the chosen model (knot points, time intervals, covariates, number of iterations, and parameters for outliers) and running the model using the `run_eg gla_lmm()` R function, several outputs are provided to aid the analyst in making a final checks; these outputs include plots of residuals, the model coefficients, the model call, a table of pairwise correlations between all derived phenotypes and a list of outliers identified for each of the latent variables. This allows the analyst to make decisions about model fit, outlier inclusion or exclusion and inform downstream analysis based on the extent of the correlation between the derived phenotypes. Finally, individual level derived phenotypes for each specified time-period are output as comma-separated files. For instance a .csv file will contain all individuals (males and females have a separate .csv file) and their slopes for infancy, childhood, late childhood, and adolescence. The csv files can then be called by the `run_eg gla_gwas()` R function or the sex stratified files can be combined and included in the analysts preferred software for GWAS.

The `run_eggla_gwas()` function performs GWAS analysis that can be implemented by any cohort with unrelated individuals (any complex data structures, such as related family structure, will need to use alternative GWAS software), where genetic dosage data is available in VCF file format. The function analyses VCF files through software packages BCFtools (v.≥1.16) (<https://github.com/samtools/BCFtools>) and PLINK2 (v.≥2.0) (<https://www.cog-genomics.org/plink/2.0>) and therefore requires access to these packages. The following actions are performed using the `run_eggla_gwas()` function: 1) format input data and perform checks for ID matching, and removing outliers on a complete case basis if required, 2) format VCFs through BCFtools and annotate if requested, 3) perform GWAS using PLINK2, 4) extract information from VCF INFO field if requested, and 5) allows for further options including: annotation files to add rsID and gene symbols

References

- 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>.
- Canouil, Mickaël, Nicole Warrington, Kimberley Burrows, and Anni Heiskala. 2024. 'Egga: Early Growth Genetics Longitudinal Analysis'. Zenodo. <https://doi.org/10.5281/zenodo.10594717>.
- 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/>.
- 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>.
- 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>.
- Järvelin, M. R., A. L. Hartikainen-Sorri, and P. Rantakallio. 1993. 'Labour Induction Policy in Hospitals of Different Levels of Specialisation'. *British Journal of Obstetrics and Gynaecology* 100 (4): 310–15. <https://doi.org/10.1111/j.1471-0528.1993.tb12971.x>.
- Laird, N. M., and J. H. Ware. 1982. 'Random-Effects Models for Longitudinal Data'. *Biometrics* 38 (4): 963–74.
- 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>.
- Perperoglou, Aris, Willi Sauerbrei, Michal Abrahamowicz, and Matthias Schmid. 2019. 'A Review of Spline Function Procedures in R'. *BMC Medical Research Methodology* 19 (1): 46. <https://doi.org/10.1186/s12874-019-0666-3>.
- Sabatti, Chiara, Susan K. Service, Anna-Liisa Hartikainen, Anneli Pouta, Samuli Ripatti, Jae Brodsky, Chris G. Jones, et al. 2009. 'Genome-Wide Association Analysis of Metabolic Traits in a Birth Cohort from a Founder Population'. *Nature Genetics* 41 (1): 35–46. <https://doi.org/10.1038/ng.271>.
- Sovio, Ulla, Amanda J. Bennett, Iona Y. Millwood, John Molitor, Paul F. O'Reilly, Nicholas J. Timpson, Marika Kaakinen, et al. 2009. 'Genetic Determinants of Height Growth Assessed Longitudinally from Infancy to Adulthood in the Northern Finland Birth Cohort 1966'. *PLoS Genetics* 5 (3): e1000409. <https://doi.org/10.1371/journal.pgen.1000409>.
- University of Oulu., 1966. 1966. 'Northern Finland Birth Cohort 1966'. 1966. <https://etsin.fairdata.fi/dataset/716939c3-7a2a-4b6a-91f3-92aca09bc52d>.
- University of Oulu., 1986. 1986. 'Northern Finland Birth Cohort 1986'. 1986. <https://etsin.fairdata.fi/dataset/f22c6599-2293-42bd-b65a-1a77945ed613>.