**Supplementary Materials**

***“Associations of arterial thickness, stiffness and blood pressure with brain morphology in early adolescence: A prospective population-based study”***

Serena Defina, MSc, Carolina C.V. Silva, PhD, Charlotte A.M. Cecil, PhD, Henning Tiemeier, PhD, Janine F. Felix, PhD, Ryan L. Mutzel, PhD, Vincent W.V. Jaddoe, PhD

Table of Contents

[Methods S1 | Imputation rationale and quality 2](#_Toc149309259)

[Methods S2 | Non-linear effects of each exposure 3](#_Toc149309260)

[Methods S3 | Linear mixed-effect model 4](#_Toc149309261)

[Methods S4 | Vertex wise cortical thickness 4](#_Toc149309262)

[Supplemental references 4](#_Toc149309263)

[Table S1 | Assessment of missing values and imputation quality 5](#_Toc149309264)

[Table S2 | Correlation matrix 6](#_Toc149309265)

[Table S3 | Main results 7](#_Toc149309266)

[Table S4 | Linear mixed model results (with age interaction) 9](#_Toc149309267)

[Table S5 | Linear mixed model results (without age interaction) 11](#_Toc149309268)

[Table S6 | Sex interaction 13](#_Toc149309269)

[Table S7 | Global vs. regional volumes 15](#_Toc149309270)

[Table S8 | Subcortical brain regions 17](#_Toc149309271)

[Table S9 | White matter tracts (FA) 19](#_Toc149309272)

[Table S10 | White matter tracts (MD) 21](#_Toc149309273)

[Table S11 | Vertex-wise analysis results 23](#_Toc149309274)

[Table S12 |Birth-related covariates 25](#_Toc149309275)

[Table S13 | Sensitivity (complete MRI data) 27](#_Toc149309276)

[Table S14 | Sensitivity (full sample) 29](#_Toc149309277)

[Figure S1 | Flowchart 31](#_Toc149309278)

[Figure S2 | DAG 32](#_Toc149309279)

[Figure S3 | Non-linear associations 33](#_Toc149309280)

[Figure S4 | Sex differences 34](#_Toc149309281)

[Figure S5 | Global vs. regional volumes 34](#_Toc149309282)

[Figure S6 | White matter tracts (FA) 35](#_Toc149309283)

[Figure S7 | White matter tracts (MD) 35](#_Toc149309284)

# Methods S1 | Imputation rationale and quality

**Imputation strategy**

Missing values in continuous variables were imputed by random forest imputation1, as implemented by the mice package (version 3.13.0), while categorical variables (i.e., sex, ethnicity, and maternal education) were imputed using standard mice method (logistic regression). We used 10 trees and 40 iterations to construct 20 imputed datasets2.

Random forest imputation is a machine learning technique, derived from classification and regression trees predictive models, which recursively subdivide the data based on values of the predictors. The technique does not rely on any distributional assumptions, and can conveniently accommodate nonlinearities and interactions, in the data, without need for them to be explicitly specified in the imputation model. While more computationally intensive compared to traditional regression-based imputation (i.e., full conditional specification), the technique is ideal for situations where, for example, multicollinearity issues, or the omission of important nonlinear terms may complicate model specification, potentially yielding biased parameter estimates and standard errors.

Records with missing values were imputed by random draws from independent normal distributions centered on conditional means predicted using random forest. The method fits each tree to a different bootstrap sample of the data (to prevent overfitting) and aggregates the results.

The auxiliary variables selected for the imputation model included: all available repeated measures of the key variables of interest, and an additional set of auxiliary variables that were selected because they are believed to be either related to missingness or to the variables of interest themselves.

These included: total brain and gray matter volumes, global fractional anisotropy and mean diffusivity, (at 10 years), total intracranial volume, cortical grey matter volume, subcortical gray matter volume, white matter volume, CSF volume, brainstem volume, cerebellum subcortical and cortical volumes (at 10 and 13 years); all individual subcortical brain region volumes and white matter tracts FA and MD (at 10 and 13 years); systolic and diastolic blood pressure (at 6 years); sex; ethnicity; age at MRI visit (both 10 and 13 years) and age at ultrasound/blood pressure measurement visit (10 years); height (at 6 and 10 years); BMI z-score (at 6 and 10 years); maternal age (at conception); maternal education (during pregnancy and when children were 3 and 6 years old); household income (at 6 years); maternal BMI (pre-pregnancy and at 6 years); gestational age and weight at birth; parity.

**Theoretical rationale**

There are two important potential consequences of missing data. The first is the decrease in *precision* (wider confidence intervals) and *power* caused by the reduction in data. The second, and more serious, is the potential for *bias* in the estimation of association parameters.

Primer: missing data mechanisms

The complexity of the missing data problem, i.e., obtaining accurate inferential estimates in the presence of missing data, depends on the nature of the mechanism by which data are missing. Little and Rubin3 provided a popular framework to describe categories of missing data mechanisms given the relationship with observed and unobserved values.

The less problematic scenario occurs when the probability of an observable data point being missing does not depend on any observed or unobserved parameters: Missing Completely at Random (MCAR). More commonly, the missingness probability depends on observed variables, and hence it can be accounted for by the information contained in the dataset: Missing at Random (MAR). The most challenging missingness mechanism occurs when the missingness probability depends on unobserved values: Missing Not at Random (MNAR).

Unfortunately, it is not possible to distinguish between MAR and MNAR mechanisms without additional external data or prior knowledge. However, the MAR assumption is usually reasonable in the context of longitudinal observational studies.

It is important to realize that the term *missing at random* does not mean that the missing data are a simple random subsample of all the data points. That scenario is MCAR. Under MAR, missing data may be more frequent in some subgroups than in others, but information defining the subgroups is observed.

Among the various available approaches to handling missing data, multiple imputation (MI) has been widely adopted and accepted by methodologists as an appropriate framework for dealing with MAR (and MCAR) mechanisms.

Although simpler solutions for handling missing outcomes are still routinely used, these approaches have been amply shown to be inadequate and even misleading, as they do not preserve important characteristics of the whole data set, such as key relationships among the variables and means. For instance, list-wise deletion (e.g., selecting on outcome availability prior to imputation) *requires MCAR data* in order to not introduce bias in the result, it makes strong assumptions about the covariance structure of the data (that it is compound symmetric), and it has been discouraged by statisticians.

In our study, we have strong reason to believe that dropout / missingness might depend at least partly on measured (or unmeasured) variables. For example, children who did have measured brain outcomes are more often female (*P*<0.001 for structural MRI; *P*=0.042 for DTI data), Dutch (*P*<0.001 in both subsamples), and first born (*P*=0.028; *P*=0.029), they come from families with higher income (*P*<0.001; *P*=0.001), they have higher gestational weight (*P*=0.001), and they have older (*P*<0.001) and more highly educated mothers (*P*=0.001).

Based on these statistics, one can speculate, for example, that families with higher socio-economic status may be more likely to bring their children to the follow-up visits. On the other end, children from low socio-economic status families would be underrepresented in the (complete-cases) sample. In other words, the MCAR assumption is likely violated.

Hence, although it is important to point out that our approach relies on MAR assumptions, and we cannot guarantee unbiased results under the situation where missingness depends on unobserved information (MNAR), these assumptions are not nearly as unrealistic as those required for a complete-case analysis. Additionally, it has been argued that MI can offer some protection against MNAR mechanisms, unlike MCAR methods.

**Missing patterns and imputation quality results**

Key variables’ descriptives before (i.e., in the original sample) and after imputation (i.e., pooled across the 20 imputed datasets) together with the number and percentage of missing values are presented in *Table S1*. Note how differences between the pooled (i.e., after imputation) and the original metrics are very small. None of these differences resulted statistically significant.

The fraction of missing values per variable ranges from 0 to a maximum of 77% (i.e., structural MRI metrics). Although the level of missingness naturally affects MI performance, we decided to leverage all available data even when missingness was extensive, as recommended by several statisticians. In support of this approach, two simulation studies also suggested that, under the condition of a large enough sample size (<1000), such levels of missingness still allow for a reasonably low expected bias if any.

Convergence

Visual inspection of the convergence graphs showed no signs of unhealthy convergence. In some cases, the trace lines showed strong initial trends and slow mixing, but regardless of the proportion of missing data, results were stable after ~20 of the 40 iterations. Convergence plots are not presented here, but are available upon request to SD.

Imputed vs. observed values

Next, we inspected and compared the density of the incomplete and imputed data. These graphs are [publicly available](https://github.com/SereDef/arterial-health-brain) for the key variables of interest. The blue line marks the observed and red lines indicate imputed values.

**Sensitivity analysis**

A thorough and sensible sensitivity analysis is an important step in producing and reporting robust estimates. Hence, we also included a sensitivity analysis to assess the extent to which our analyses were robust to missing data assumptions. None of the main conclusions were impacted.

# Methods S2 | Non-linear effects of each exposure

To assess potential non-linearities in the relationship between each arterial health marker (i.e., IMT, Distensibility, SBP and DBP) and the brain outcomes or interest (i.e., TBV, GMV, FA and MD), we tested a set of 8 non-linear models (i.e., one for each exposure – outcome pair).

More specifically, we first determined whether evidence of a non-linear association was present based on global F-test(s) comparing the linear models to four polynomial alternatives (i.e., 2nd- to 5th- degree polynomials). We considered non-linearity present if at least one of these comparisons was statistically significant (P < 0.05).

To then characterize the shape of these non-linear associations, we plotted natural spline terms using 4 knots. This approach allowed us to adequately represent the complexity of these relationships, without incurring overfitting problems due to unduly complex models.

# Methods S3 | Linear mixed-effect model

To assess the relationship between arterial health and brain changes from 10 to 14 years (i.e., two timepoints), we ran a linear mixed-effect effect model using the lmer() function from the lme4 R package.4

Models included a random intercept and a fixed-effect part corresponding to the interaction between each exposure of interest (i.e., IMT, distensibility, SBP and DBP) and age at MRI measurement + the base and full set of covariates. When the interaction terms between each exposure and age resulted not statistically significant, they were removed from the models, to grant more interpretable main effects of each exposure. The optimization criterium was set to maximum likelihood (i.e., REML=FALSE).

Note that, outcomes were included in the model in their original scale (cm3), rather than their standardized version, because centering dependent variables in mixed-effects models which include random effects per subject has been shown to result in incorrect estimated correlations, likelihood values and predictions.4

P-values for fixed effects were computed via the Satterthwaite’s degrees of freedom approximation, as implemented by the R package lmerTest.

# Methods S4 | Vertex wise cortical thickness

Exploratory vertex-wise cortical thickness analyses were performed in R using the [QDECR](https://github.com/slamballais/QDECR) package.

Cortical thickness was calculated as the closest distance from the grey-white matter boundary to the grey-cerebrospinal fluid boundary at each vertex on the tessellated surface. The labeling of clusters identified in the brain was performed using the technique of Klein & Tourville. The cluster-forming threshold was set to *p*=0.001, as it has shown to correspond closely to a false-positive rate of 0.05. Prior to analyses images were smoothed using a 10 mm full-width-at-half-maximum Gaussian kernel. Reconstructed images were visually inspected and images that were rated as poor were excluded.

# Supplemental references

1. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. *Am J Epidemiol*. Mar 15 2014;179(6):764-74. doi:10.1093/aje/kwt312

2. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 12/12 2011;45(3):1 - 67. doi:10.18637/jss.v045.i03

3. Rubin DB. Multiple imputation for survey nonresponse. New York: Wiley; 1987.

4. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 10/07 2015;67(1):1 - 48. doi:10.18637/jss.v067.i01

# Table S1 | Assessment of missing values and imputation quality

# Table S2 | Correlation matrix

# Table S3 | Main results

# Table S4 | Linear mixed model results (with age interaction)

# Table S5 | Linear mixed model results (without age interaction)

# Table S6 | Sex interaction

# Table S7 | Global vs. regional volumes

# Table S8 | Subcortical brain regions

# Table S9 | White matter tracts (FA)

# Table S10 | White matter tracts (MD)

# Table S11 | Vertex-wise analysis results

# Table S12 |Birth-related covariates

# Table S13 | Sensitivity (complete MRI data)

# Table S14 | Sensitivity (full sample)

# Figure S1 | Flowchart

Flowchart detailing exclusion criteria. Prior to analysis, all twins were excluded, and only one non-twin sibling was retained in the sample (based on data completeness, or, if that was equal between siblings, randomly).

# Figure S2 | DAG

Directed acyclic graph (DAG) representing the assumed causal structure underlying the observed data. Variables which were included in the models as covariates are indicated by a black square box.

# Figure S3 | Non-linear associations

For each exposure - outcome pair (**A-D.** total brain volume (TBV); **E-H.** grey matter volume (GMV); **I-L.** global fractional anisotropy (FA); and **M-P.** mean diffusivity (MD) the linear and non-linear fits are presented. Linear associations (black dashed line) were pooled across datasets, while non-linear associations (green continuous lines) were fit in each imputed dataset individually using natural splines with 4 knots. The gray shadows also mark the –2.5 and +2.5 SD cutoffs of exposure distribution.

# Figure S4 | Sex differences

For each exposure of interest (**A.** intima-media thickness (IMT); **B.** carotid distensibility; **C.** Systolic blood pressure (SBP); and **D.** Diastolic blood pressure (DBP)), the standardized association estimates and their 95% confidence intervals are displayed on the x-axis for each outcome (total brain volume (TBV); grey matter volume (GMV); global fractional anisotropy (FA); and mean diffusivity (MD)), in light blue for boys and pink for girls. The corresponding FDR-corrected P-values are also reported.

# Figure S5 | Global vs. regional volumes

Sankey-like plot representing associations between SBP (above) / DBP (below) and each morphological (i.e., volumetric) brain marker: total intracranial volume (TIV), total brain volume (TBV), cerebro-spinal fluid (CSF), gray and white matter volumes, ventricular volume, subcortical and cortical grey matter volumes and each of the seven subcortical brain regions. The height of each square represents the magnitude of the corresponding estimate and the color indicates its direction (pink for positive associations and blue for negative ones). Estimates that were statistically significant (after adjustment for all covariates and FDR correction) are indicated by an \*. Estimates marked with ° were nominally significant, but did not survive FDR correction. Structural relationships between these brain outcomes are represented by the grey lines, with increasing regional specificity from left to right.

# Figure S6 | White matter tracts (FA)

For each exposure of interest (**A.** intima-media thickness (IMT); **B.** carotid distensibility; **C.** Systolic blood pressure (SBP); and **D.** Diastolic blood pressure (DBP)), the standardized association estimates and their 95% confidence intervals are displayed on the x-axis for each white matter tract (Cingulate gyrus, Cortico-spinal tract, Uncinate fasciculus, Inferior & Superior longitudinal fasciculus, Major & Minor forceps) average FA value. The corresponding FDR-corrected P-values are also reported.

# Figure S7 | White matter tracts (MD)

For each exposure of interest (**A.** intima-media thickness (IMT); **B.** carotid distensibility; **C.** Systolic blood pressure (SBP); and **D.** Diastolic blood pressure (DBP)), the standardized association estimates and their 95% confidence intervals are displayed on the x-axis for each white matter tract (Cingulate gyrus, Cortico-spinal tract, Uncinate fasciculus, Inferior & Superior longitudinal fasciculus, Major & Minor forceps) average MD value. The corresponding FDR-corrected P-values are also reported.