Supplemental Online Content

Koc D, Tiemeier H, Stricker BH, Muetzel RL, Hillegers M, El Marroun H. Prenatal antidepressant exposure and offspring brain morphologic trajectory. *JAMA Psychiatry*. Published online August 30, 2023. doi:10.1001/jamapsychiatry.2023.3161

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods. Neuroimaging, Maternal Depressive Symptoms, Psychopathology and SSRI Use, Pregnancy Complications, Covariates, Statistical Analyses

Maternal depressive symptoms, psychopathology score and SSRI use

The score for psychopathology was determined using the Global Severity Index (GSI) (range, 0-4, with higher scores indicating more clinically relevant psychological symptoms), which was derived from the BSI.[1] We found a strong correlation between the prenatal GSI score and the prenatal BSI score for depression (Spearman correlation coefficient = 0.8). The levels of depressive symptoms were found to have a moderate correlation across all of the assessments (Spearman correlation coefficients ranging from 0.47 to 0.59), with the strongest correlation existing between the levels of depressive symptoms measured at 20 weeks of gestation and those measured when the child was 2 months old.

In a subgroup of 905 women from the Generation R cohort, we evaluated the performance of the Brief Symptom Inventory (BSI) in detecting clinical depression using the recommended cut-off score.[2] To assess clinical depression in this subgroup, we used the Composite International Diagnostic Interview (CIDI), which is a structured interview based on DSM-IV criteria and has been reported to have good reliability and validity.[3] We calculated the positive likelihood ratio (LR+) as it is more suitable for detecting low prevalence conditions, and the calculated LR+ was 7.29.[2] This indicates that the cut-off score has moderate diagnostic accuracy in identifying clinical depression.

Out of the 41 pregnant women who reported using SSRIs during pregnancy, 20 women used SSRIs during the first trimester only, 21 women used SSRIs in the first trimester and in one or two additional trimesters. The SSRIs used (n, median mg/day (range: min-max) included paroxetine [22, 20 mg/day (20-40)], fluoxetine [12, 20 mg/day (20-40)], sertraline [8, 50 mg/day

(50-100)], fluvoxamine [4, 50 mg/day (50-100)], and citalopram [4, 20 mg/day (20-20)].The total number of SSRIs reported exceeded the number of participants as some pregnant women used multiple SSRIs. The mean duration of SSRI use was approximately 8 months (250 days). The reasons for using SSRIs were mainly due to (previous) depression (n=30) or (previous) anxiety (n=3), whereas the reasons were unknown for 8 cases.

Out of the 77 women who reported SSRI use prior to pregnancy, the most frequently used SSRIs (n, median mg/day (range: min-max) were paroxetine [43, 20 mg/day (20-40)], followed by citalopram [13, 20 mg/day (20-20)], sertraline [9, 50 mg/day (50-100)], fluoxetine [10, 20 mg/day (20-40)], and fluvoxamine [2, 50 mg/day (50-50)].All women who reported SSRI use prior to pregnancy had discontinued their medication at least 3 months before the start of pregnancy. The average length of SSRI use was 150 days (~5 months).

Only 12 out of 41 of the SSRI-using pregnant women reported clinically relevant depressive symptoms, and 14 of 77 women using SSRI before pregnancy only reported clinically relevant depressive symptoms.

Pregnancy complications

We assessed multiple pregnancy complications, including gestational hypertension, preeclampsia, preterm birth, low birth weight, and low Apgar score. We defined gestational hypertensive disorders, which encompassed preeclampsia and gestational hypertension, based on the criteria established by the International Society for the Study of Hypertension in Pregnancy.[4] Preterm birth was classified as a delivery that occurred before 37 weeks of gestation, and low birth weight was defined as a birth weight of less than 2,500 g. A low Apgar score was defined as a score of less than 7 at the 5-minutes.

Covariates

Maternal national origin was categorized as Dutch, non-Dutch Western and non-Dutch non-Western according to the classification of Statistics Netherlands. Maternal education was divided into three categories: primary or lower (no education/primary school), secondary (high school/vocational training), and higher (higher vocational training/university) and household income in pregnancy were categorized into less than €1200 (poverty), €1200 to €2000 (low income in 2005), and more than €2000 per month. In each trimester, questionnaires were used to collect information on maternal tobacco use, which was categorized into three categories: never, until pregnancy was known, and continued during pregnancy while alcohol use was categorized into four categories: never, until pregnancy was known, continued to drink occasionally during pregnancy and continued to drink frequently during pregnancy.

Image acquisition and neuroimaging processing

All children were familiarized with neuroimaging using a mock session. At T1, images were acquired using an inversion recovery fast spoiled gradient recalled (IR-FSPGR) sequence (sequence parameters: TE: 4.2ms, TR: 10.3ms, TI: 350ms, flip angle: 16°, acquisition time: 5 min 40 sec, FOV: 230.4x230.4, in-plane resolution: 0.9mm3, coverage: whole-brain)[5]. At T2 and T3, images were acquired using a 3D coronal inversion recovery fast spoiled gradient recalled (IR-FSPGR, BRAVO) sequence (sequence parameters: TE: 3.4ms, TR: 8.77ms, TI: 600ms, flip angle: 10°, acquisition time: 5 min 20 sec, FOV: 220x220, in-plane resolution: 1.0 mm^3, phase encoding: R/L, fat suppression: yes, coverage: whole-brain).[6] For both scanners, an 8-channel receive-only head coil was used. T_1-weighted structural images were acquired with an inversion recovery–prepared fast spoiled gradient recalled sequence.

The quality of FreeSurfer output was visually inspected by at least one rater. Each scan was scored on a Likert scale based on how successfully FreeSurfer defined the gray-white matter and

outer gray matter boundaries. Raters included trained researchers, and the task was completed after they accurately rated 30 images whose quality had already been verified. Scan results at T1 and T2 were rated on a five-point Likert scale (unusable, poor, sufficient, good, excellent). Scan results at T3 were rated on a three-point Likert scale (poor, questionable, good) by two independent raters. All scans that were unusable or poor quality reconstruction were excluded from further analyses.[7]

Metrics of volume, including total brain volume, total grey and white matter volume, volumes of the specific limbic structures (hippocampus and amygdala) and cortical brain structures including volumes of frontal, cingulate, somatosensory (postcentral gyrus) and higher-order visual areas were extracted. Left and right volumes were averaged we had no a priori hypotheses on lateralized association of prenatal SSRI exposure and brain morphometry.

Statistical Analyses

Differences in the demographic data were tested using analysis of variance (ANOVA), chi-square test and the Kruskal-Wallis test.

eFigure 3 outlines the analytic considerations for modelling strategy used in this study. While prenatal effects are essential and of great interest, it is also clear that postnatal effects (particularly the early postnatal period) have a meaningful influence on offspring's neurodevelopment. Therefore, one must also rule out whether these effects are based on postnatal maternal depressive symptoms experienced by offspring. According to the scenario in eFigure 3A, postnatal maternal depressive symptoms is a variable that lies along the path as a mediator between the exposure (A) and the outcome (Y). Controlling for a mediator will hinder an effort to determine the combined effect of predictors (A → Y plus A → M → Y) on the

outcome by blocking one proposed causal pathway. To avoid inappropriate control of mediator, we adopted an additional exposure group that includes postnatal depressive symptoms exposure only (B → Y) (eFigure 3B).

The equation below outlines the general modelling strategy used in basic model.

$$Linear\ model$$

$$Cerebral\ Volume_{ij} = intercept + \beta_1 \times Group_i + \beta_2 \times (cAge\_c_{ij})$$

$$+ \beta_3 \times (Group \times cAge\_c_{ij}) + \beta_4 \times Sex_i + e_{ij}$$

$$Quadratic\ model$$

$$Cerebral\ Volume_{ij} = intercept + \beta_1 \times Group_i + \beta_2 \times (cAge\_c_{ij}) + \beta_3 \times (cAge\_c_{ij})^2$$

$$+ \beta_4 \times (Group \times cAge\_c_{ij}) + \beta_5 \times (Group \times (cAge\_c_{ij})^2) + \beta_4 \times Sex_i + e_{ij}$$

In the equation, i is an index of subject, j is an index of the measurement and e_{ij} is the error term. The intercept and β terms were considered as fixed effects. To allow for the interpretation of intercept differences, maternal age (mAge_c = 30.9 years) and child age (cAge_c = 11.4 years) were centered at the sample's mean age. For each structure, linear and quadratic developmental trajectories were modelled. To determine whether a quadratic or linear growth model best fit the association, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to evaluate the most suitable model across different degrees of polynomial fit.[8,9] After determining the best growth model for the entire sample, we used likelihood-ratio tests (LRT) to determine if either the growth curve height or shape (combined exposure effect) were statistically different between each exposure group and reference. For each structure, we compared the best fitting growth model (as described above) including the

exposure group (i.e. prenatal SSRI exposure vs reference or pre- and postnatal depression exposure vs reference) against the same model without exposure groups.

Since we have an unbalanced design in the sense that the number of observations is not constant across exposure and reference groups, the choice of method for variance component estimation and approximation of the degrees of freedom is critical. To reduce bias in the estimated variance-covariance matrix of fixed effects, models are fitted using restricted maximum likelihood (REML) and *p*-values are derived using Kenward-Roger approximation that produces acceptable Type 1 error rates for smaller samples.[10,11]

We used sensitivity analyses to compare included with those in the overall sample to assess the robustness and reliability of our findings. In these sensitivity analyses, we applied inverse probability of attrition weighting (IPAW) model to account for potential selection bias.[12,13] Weights were estimated using logistic regression models for each individual so that results would be representative for the initial population of this cohort study. Following, we used these estimated weights to our linear mixed models to compared unweighted results with IPAW. Post-hoc group analyses were performed by defining a subgroup of children exposed to prenatal SSRI but not to clinically relevant depressive symptoms (n=29 with 56 scans) in the analyses of primary brain outcome measures (i.e. 12 children exposed to SSRI and clinical depressive symptoms were excluded.

To determine the impact of influential data points, we used the 'influence.ME' package[14] and calculated a measure of influence which also accounts for the nested data structure. The influence of individual observations on single parameter estimates was assessed by DFBETAS. A cut-off value is given by DFBETAS[14], which is $2/\sqrt{n}$, where n is the number of observation (n=5624). We performed a test for changes in the statistical significance of fixed parameter

estimates for estimates when DFBETAS values exceeded this cut-off value (<-0.025 or >0.025). Values exceeding this cut-off value may overly influence the regression outcomes for that specific estimate. We specify the t-value as a test for significance at a commonly used value (<-1.96 or >1.96).[14]

All analyses were conducted with the R statistical software version 4.1.2 [15] including the lme4-package for the longitudinal mixed effects analyses with unweighted and IPAW-weighted models and other packages for statistics testing (e.g. broom.mixed, emmeans, p.adjust, and ipw-package) and visualization (e.g. ggplot2, ggseg, and sjPlot package) of the results. A false discovery rate correction (FDR-Benjamini Hochberg) was applied separately within primary and secondary outcomes in all tests to control for type I errors[16] (i.e., global and subcortical analyses [4 brain structures], cortical brain areas analyses [8 brain structures]).

eResults. Nonresponse and Sensitivity Analyses

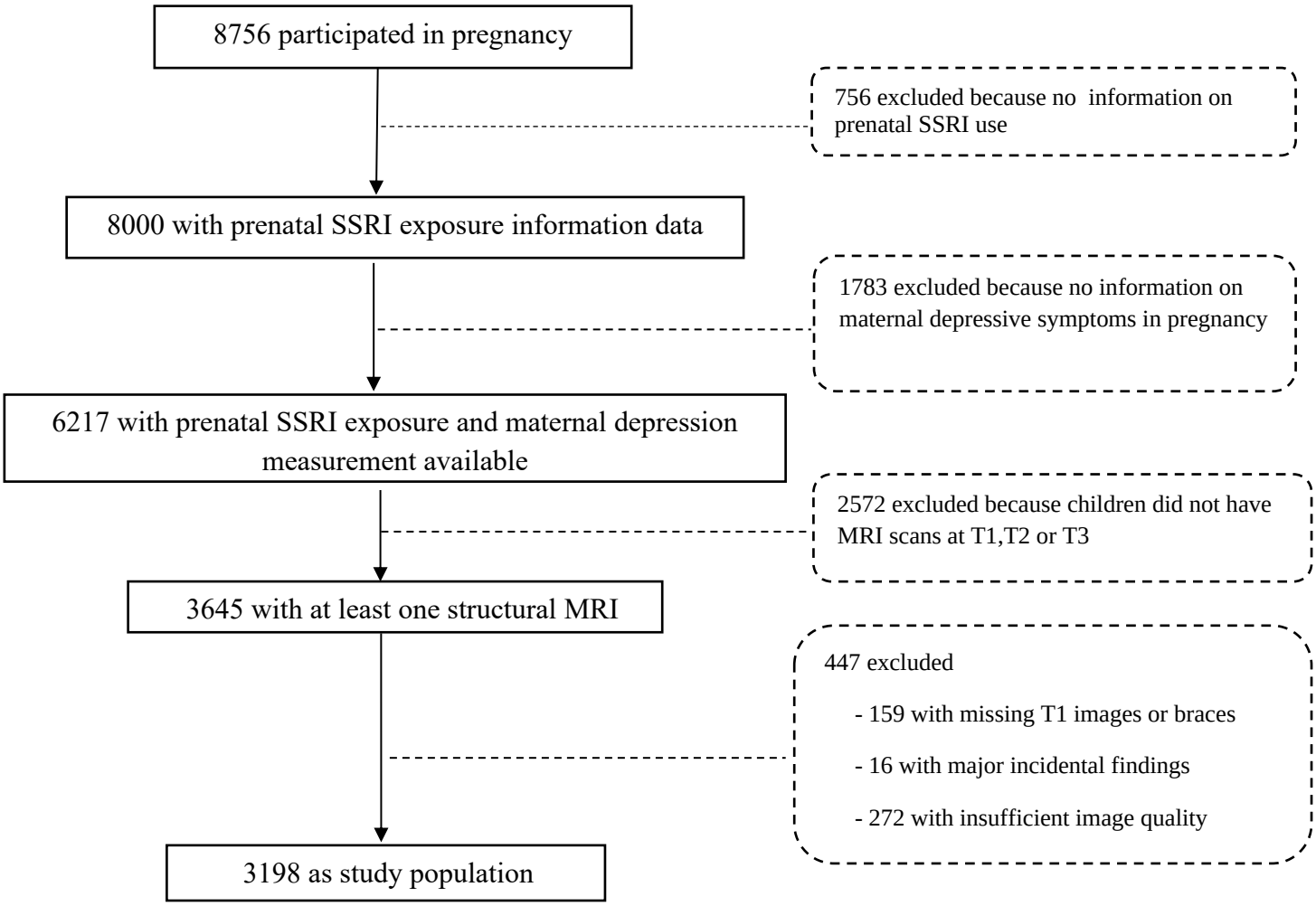
We compared the characteristics of the 3198 mother-child dyads included in the analyses with at least one neuroimaging data to those of 3019 mother-child dyads who were lost to follow-up (eTable 3). Compared to responders, the non-responders were younger (29.2 years (s.d. = 5.3) than responders (30.9 years (s.d. = 4.7), t = -13.8, $p < .001$), were more likely to be of non-Dutch origin (47.1% v. 60.7% Dutch, $\chi^2 = 113.5$, $p < .001$), less educated (36.9% v. 51.6% higher education; $\chi^2 = 158.2$, $p < .001$), smoked tobacco more often in pregnancy (66% v. 72.9% never smoked in pregnancy, $\chi^2 = 46.4$, $p < .001$) and consume less alcohol (50.6% v. 38.1% never drink alcohol in pregnancy, $\chi^2 = 99.9$, $p < .001$). The results of IPAW-weighted models did not improve the non-weighted models (eTable 5).

To further address the association of maternal depression and SSRI exposure on brain development, we conducted sub-group analyses based on the presence of depressive symptoms in the SSRI exposure group. Prenatal SSRI exposure without depressive symptoms (n=29 with 56 scans) was associated with less cerebral grey matter volume (β = -18766.4 mm^3, SE = 8567.4, p = .01), white matter (β = -13421.8 mm^3, SE = 5211.2, p = .003) and amygdala volume (β = -124.4 mm^3, SE = 34.4, p = .003) compared to the reference. The results generally remained in these post-hoc analyses, albeit with some minor attenuation compared to analyses that did not exclude children (n=12) with exposure to both SSRI and clinical depressive symptoms.

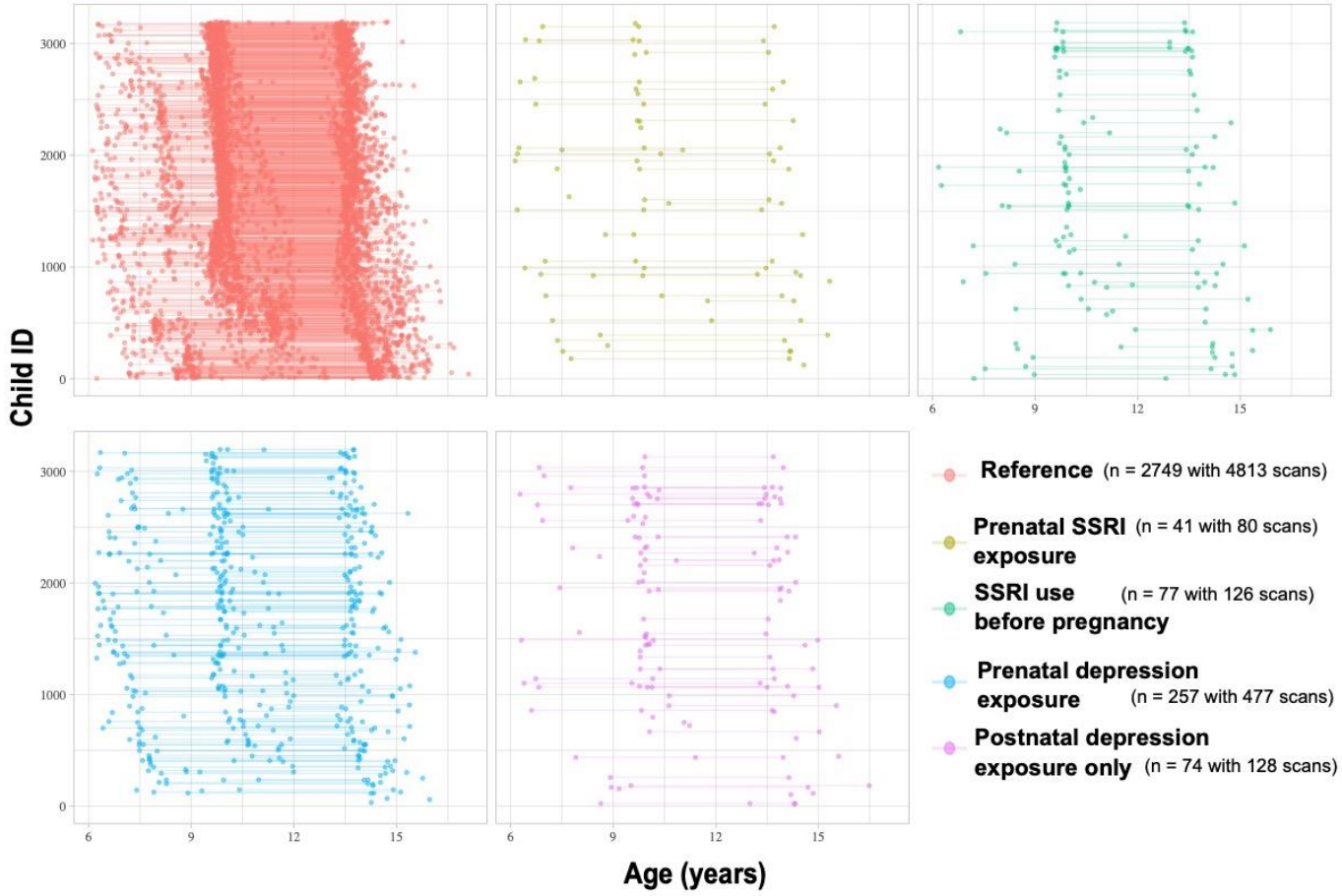
Additionally, we performed analyses to detect potentially influential observations for primary outcomes (eFigure 5). A test for changes in the statistical significance of fixed parameter estimates was applied when DFBETAS values exceeded the cut-off value (<-0.025 or >0.025) (eFigure 5).

In eTable 6, the first column “Altered Test Statistic” shows the value of the test statistic after the omitting of the potentially influential observations (indicated in the row labels). The “Changed Significance” column indicates whether the level of significance changed after each of the observations was deleted. We did not observe any changes in significance for primary outcomes (eTable 6).

eFigure 1. Flow Diagram of Study Population

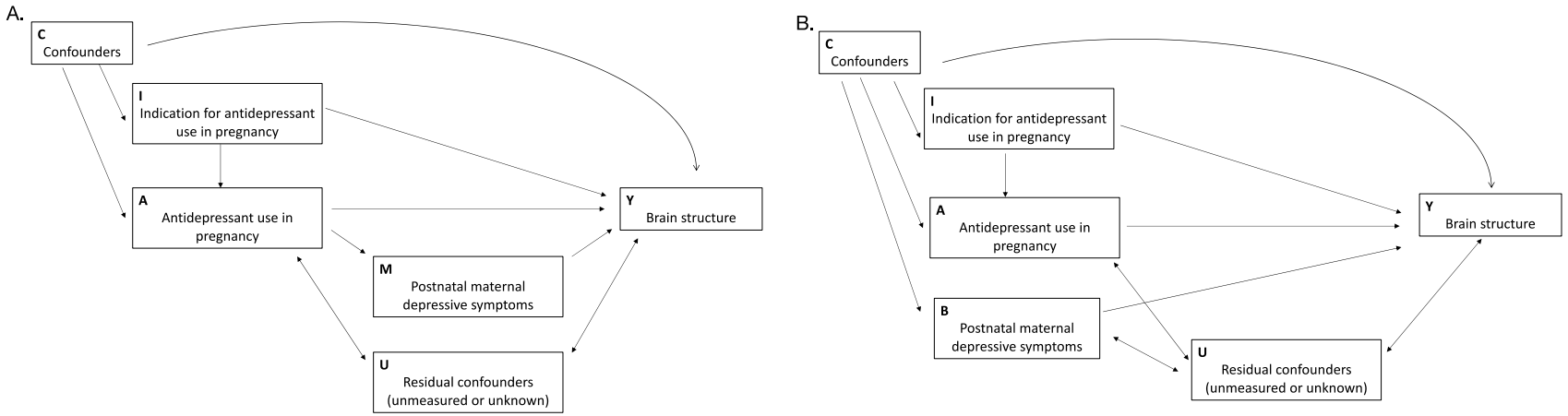


eFigure 2. Showing the Age of Each Participant at Each Study Time Point

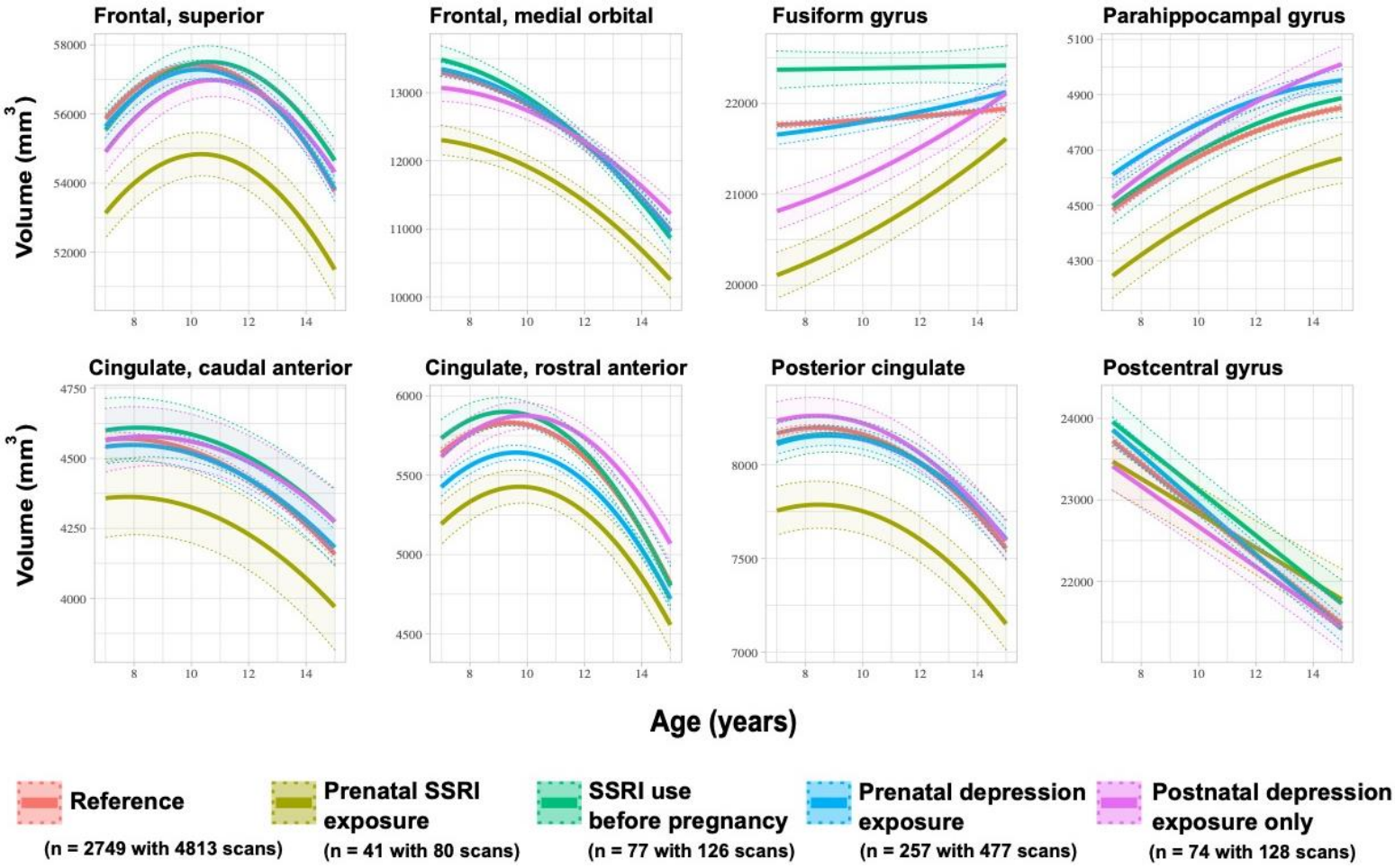


Note: Each of the 5624 obtained scans is represented by a circle, and each of the 3198 subjects is represented by a distinct row, with their scans connected by a straight line.

eFigure 3. Direct Acyclic Graph Showing the Hypothesized Relationship Between Maternal SSRI Use During Pregnancy and Brain Outcomes

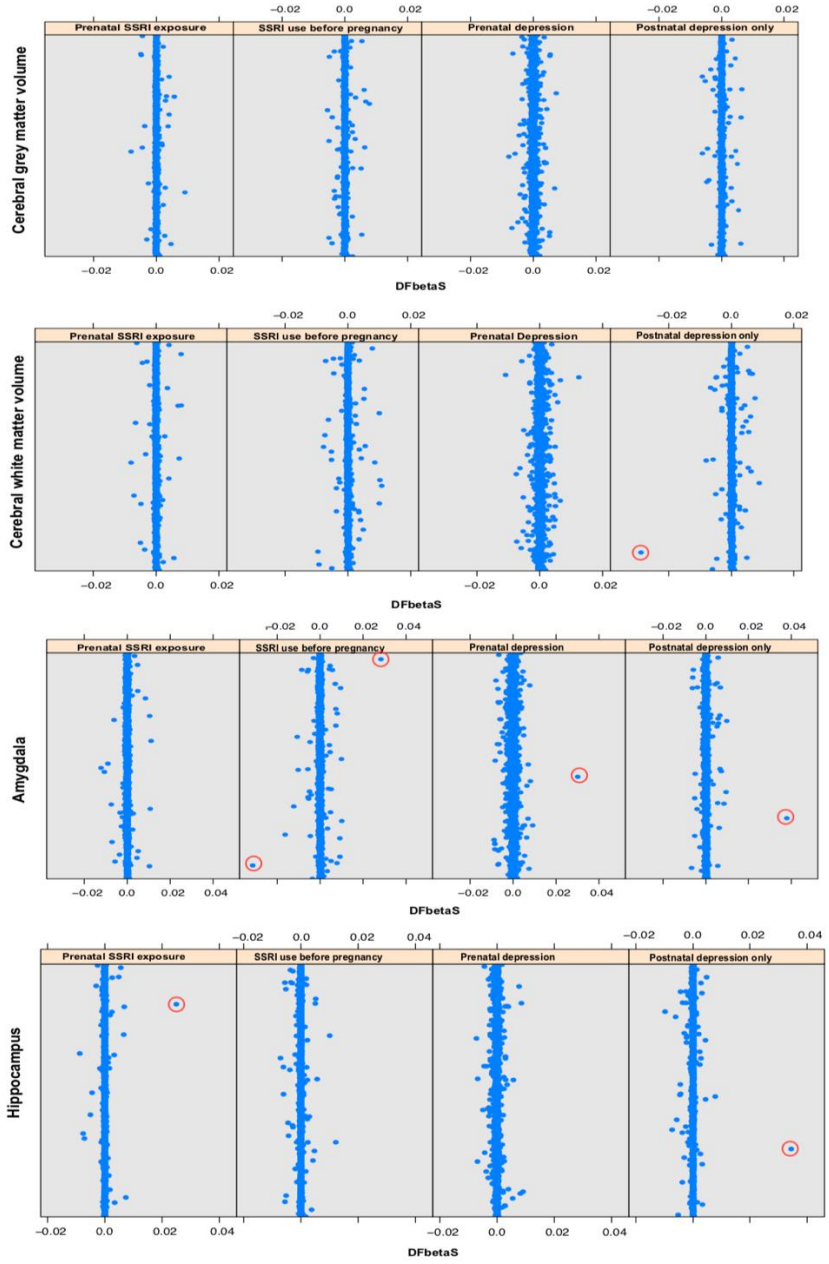


eFigure 4. Developmental Trajectories of Cortical Brain Structural Volumes



Note: Estimated average trajectories are shown with 95% confidence intervals for cortical brain volumes for each exposure group including the reference separately, ages 7 to 15 years. The adjustment was made for child sex and age at the neuroimaging assessment, maternal age at intake, maternal national origin, marital status, maternal education level, maternal substance use (tobacco, cannabis, alcohol) benzodiazepine usage during pregnancy, monthly household income and ICV.

eFigure 5. Difference in Parameter Estimates (DFBETAS) for Each Data Point and All Exposure Groups in the Analyses of Primary Outcomes



Note: To identify potential influential observations in the analyses of the primary outcomes, including cerebral gray and white matter volume, amygdala, and hippocampus volume, we performed model diagnostic analyses. We examined the distribution of DFBETAS, a measure of the influence of each observation on the regression coefficient, for each exposure group. Observations that had a DFBETAS value that exceeded a cut-off value of < -0.025 or > 0.025 were identified as potentially influential and marked with a red circle. We determined whether their exclusion from the analysis would have a significant impact on the primary outcomes. The results of these analyses are described in eTable 6.

eTable 1. Number of Images of the Participants in Each Group at the Three Assessments

Time point	Reference **(n = 2749)**	Prenatal SSRI exposure **(n = 41)**	SSRI use before pregnancy **(n = 77)**	Prenatal depression exposure **(n = 257)**	Postnatal depression exposure only **(n = 74)**	Total number of images
T1	647	24	20	119	20	830
T2	2199	28	58	182	55	2522
T3	1967	28	48	176	53	2272
All time points	**4813**	**80**	**126**	**477**	**128**	**5624**

Note: The bold numbers refer to unique participants, the non-bold numbers refer to the brain images. SSRIs, selective serotonin reuptake inhibitors. Reference group: no SSRI use, a low score on depression symptoms during pregnancy; Prenatal SSRI exposure: children exposed to SSRIs during pregnancy; SSRI use before pregnancy, a low score on depression symptoms during pregnancy: mother use SSRI before the pregnancy; Prenatal depression exposure: children exposed to clinically relevant depressive symptoms during the pregnancy; Postnatal depression exposure only: no SSRI use, children exposed to clinically relevant depressive symptoms only in postnatal period.

eTable 2. Comparison of Linear and Nonlinear (Quadratic) Models Using the Likelihood Ratio Test

Brain morphometry	Linear model		Quadratic model		
	AIC	**BIC**	**AIC**	**BIC**	**P-value**
Global and subcortical volumes					
Cerebral gray matter volume	139303	139330	139208	139242	< .001
Cerebral white matter volume	133860	133887	133422	133455	< .001
Amygdala[a]	83947	83974	83682	83715	< .001
Hippocampus[a]	88031	88057	88032	88065	.39
Cortical brain volumes[a]					
Frontal, superior	113935	113746	113746	113780	< .001
Frontal, medial orbital	100600	100732	100566	100705	< .001
Fusiform	102809	102835	102789	102822	< .001
Parahippocampal gyrus	102682	102709	102638	102716	< .001
Cingulate, caudal anterior	89607	89633	89568	89601	< .001
Cingulate, rostral anterior	96104	96131	95958	95991	< .001
Posterior cingulate	108813	108839	108659	108692	< .001
Postcentral gyrus	102137	102270	102137	102276	.126

Note: In all children (n=3198), non-linear models with a quadratic term were compared to linear models using the likelihood ratio test. A p-value <.05 suggests an improved model fit of the non-linear models. A fully adjusted model, corrected for child sex and age at the neuroimaging assessment, maternal age at intake, maternal national origin, marital status, maternal education level, maternal substance use (tobacco, cannabis, alcohol), benzodiazepine usage during pregnancy and monthly household income. AIC = Akaike information criterion. BIC = Bayesian information criterion.

[a]Intracranial volume was additionally adjusted for.

eTable 3. Nonresponse Analysis

Characteristics	Responders[a] (n = 3198)	Non-responders (n= 3019)	P-value[b]
Child sex, male (%)	47.8	51.4	.004
Gestational age at birth, week	39.8 (1.7)	39.7 (1.9)	.03
Maternal national origin (%)			
Dutch	60.7	47.1	<.001
Non-Dutch Western	8.4	9.7	
Non-Dutch Non-Western	30.9	43.2	
Marital status, with partner (%)	88.7	85	<.001
Education level			
Primary or lower	6.1	12.1	<.001
Secondary	41.7	49.4	
Higher	52.2	38.5	
Household income in €/month (%)			
<1200	15.1	25.3	<.001
1200–2000	14.4	15.5	
>2000	70.6	59.2	
Maternal tobacco use (%)			
Never during pregnancy	76.9	69.9	<.001
Until pregnancy was known	8.5	8.8	
Continued during pregnancy	14.6	21.2	
Maternal cannabis use (%)			
Never during pregnancy	94	93.6	.41
Before pregnancy	3.1	3.4	
During pregnancy	2.9	3	
Maternal alcohol use (%)			
Never during pregnancy	38.1	50.6	<.001
Until pregnancy was known	14	12.7	
Throughout pregnancy, occasionally	37.4	30.2	
Throughout pregnancy, frequently	10.4	6.5	
Used benzodiazepine in pregnancy %	2.3	3.5	.01
Maternal age at intake, mean (s.d.) (years)	30.95 (4.75)	29.21 (5.3)	<.001
Maternal Global Severity Index (GSI) (20 weeks of gestation), mean score (s.d.)	0.27 (0.3)	0.32 (0.4)	<.001
Maternal depressive symptom (BSI), mean score (s.d.)			
20 weeks of gestation	0.22 (0.4)	0.27 (0.5)	<.001
Child age 2 months	0.21 (0.4)	0.31 (0.6)	<.001
Child age 6 months	0.24 (0.5)	0.30 (0.6)	<.001

Note: Non-respondents are participants with data on maternal depression and prenatal SSRI exposure at baseline, but no at least useable neuroimaging data at follow-up.

[a]Imputed data were reported (except for maternal depressive symptom and GSI scores).

[b]P-values were derived from t-tests or Wilcoxon tests for continuous variables and chi-square tests for categorical variables.

eTable 4. Association of Prenatal SSRI Use and Maternal Depressive Symptoms With Cortical Brain Volumes

	Frontal, superior		Frontal, medial orbital		Fusiform gyrus		Parahippocampal gyrus	
	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
Prenatal SSRI exposure > Reference								
Group	-2541.7 (585.6)	.001[a]	-880.9 (220.4)	.005[a]	-1112.1 (221.3)	< .001[a]	-215.2 (67.9)	.005[a]
Group*Age	638.3 (1553.5)	.68	74.2 (67.2)	.26	168.3 (51.4)	.003[a]	13.7 (28.4)	.61
Combined exposure effect	LRT: 10.9	.004[a]	LRT: 12.7	< .001[a]	LRT: 15.7	< .001[a]	LRT: 7.1	.03
SSRI use before pregnancy > Reference								
Group	-174.5 (1188.8)	.81	70.8 (149.9)	.62	562.5 (453.2)	.13	23.5 (80.2)	.86
Group*Age	167.9 (169.7)	.32	-32.8 (65.8)	.60	-12.3 (57.6)	.82	4.2 (12.5)	.75
Combined exposure effect	LRT: 1.3	.51	LRT: 0.4	.83	LRT: 0.9	.66	LRT: 0.6	.71
Prenatal depression exposure >Reference								
Group	-106.2 (602.5)	.63	23.3 (114.2)	.73	12.9 (34.8)	.41	116.1 (63.4)	.09
Group*Age	509.1 (767.4)	.50	23.2 (31.1)	.44	40.4 (27.3)	.14	1.9 (6.3)	.82
Combined exposure effect	LRT: 4.4	.09	LRT: 0.01	.99	LRT: 1.7	.44	LRT: 4.8	.06
Postnatal depression exposure only >Reference								
Group	-341.3 (910.8)	.81	32.8 (145.9)	.62	-480.5 (189.2)	.002[a]	90.5 (80.2)	.86
Group*Age	264.9 (154.7)	.08	74.8 (60.8)	.21	140.3 (51.6)	.003[a]	13.2 (12.5)	.17
Combined exposure effect	LRT: 1.3	.51	LRT: 0.4	.83	LRT: 9.4	.006[a]	LRT: 0.6	.71
	Cingulate, caudal anterior		**Cingulate, rostral anterior**		**Posterior cingulate**		**Postcentral gyrus**	
	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
Prenatal SSRI exposure > Reference								
Group	-200.2 (95.6)	.04	-371.7 (94.2)	< .001[a]	-409.1 (110.8)	.003[a]	26.2 (44.9)	.30
Group*Age	19.7 (18.5)	.52	76.4 (43.2)	.07	29.3 (16.4)	.10	69.7 (53.4)	.19
Overall exposure effect	LRT: 2.3	.22	LRT: 10.9	.006[a]	LRT: 6.5	.03	LRT: 0.7	.70
SSRI use before pregnancy > Reference								
Group	57.5 (260.8)	.54	49.8 (249.9)	.86	9.5 (138.2)	.70	244.5 (668.2)	.74
Group*Age	11.8 (19.8)	.44	-7.8 (41.8)	.84	13.3 (16.6)	.44	2.3 (53.5)	.96
Combined exposure effect	LRT: 2.1	.35	LRT: 0.2	.90	LRT: 1.5	.46	LRT: 1.6	.44

Prenatal depression exposure >Reference								
Group	-5.5 (15.5)	.54	-166.3 (65.1)	.006[a]	4.9 (14.8)	.06	39.1 (163.4)	.31
Group*Age	14.7 (8.4)	.09	52.2 (80.1)	.38	27.2 (80.2)	.70	-24.1 (25.3)	.33
Combined exposure effect	LRT: 1.5	.47	LRT: 9.1	.006[a]	LRT: 5.2	.04	LRT: 3.1	.20
Postnatal depression exposure only > Reference								
Group	46.5 (160.8)	.79	89.8 (223.4)	.86	56.5 (132.2)	.74	-173.5 (335.2)	.69
Group*Age	20.8(16.8)	.22	50.8 (38.8)	.19	6.3 (16.6)	.70	34.2 (47.5)	.47
Combined exposure effect	LRT: 1.1	.32	LRT: 0.9	.85	LRT: 2.6	.32	LRT: 1.6	.44

Note: Linear mixed-effect models were used to test the associations of prenatal SSRI and maternal depressive symptoms exposure and repeated brain morphological outcomes. Table presents effect estimates including the main effect (size) and interaction effect estimates (β) (volume change; interaction of age/-squared*group), as well as standard errors (SE) and *p* values. Main effect estimates the difference in volume (intercept, in mm^3) in the exposure groups versus the reference group. For determine the combined exposure effect on growth curve height and shape, the likelihood ratio tests (LRT) were performed for each structure. A fully adjusted model, corrected for child sex and age at the neuroimaging assessment, maternal age at intake, maternal national origin, marital status, maternal education level, maternal substance use (tobacco, cannabis, alcohol), benzodiazepine usage during pregnancy, monthly household income. ICV was additionally adjusted for subcortical volumes.

[a] Indicates significant associations after FDR correction for multiple testing.

eTable 5. Comparison of Results From the Unweighted Model With Inverse Probability of Attrition Weighting (IPAW) Models for the Associations of Prenatal SSRI and Maternal Depressive Symptoms Exposure and Brain Morphology

Brain volumes	Unweighted model		IPAW-weighted model		
	AIC	BIC	AIC	BIC	P-value
Global and subcortical volumes					
Cerebral gray matter volume	138094	138253	138209	138368	< .001
Cerebral white matter volume	133272	133431	133469	133622	< .001
Amygdala[a]	83429	83582	83507	83659	< .001
Hippocampus[a]	87549	87702	87571	87571	< .001
Cortical volumes[a]					
Frontal, superior	113297	113450	113065	113224	< .001
Frontal, medial orbital	100572	100731	100606	100758	< .001
Fusiform gyrus	102081	102233	102235	102388	< .001
Parahippocampal gyrus	86021	86174	85947	86106	< .001
Cingulate, caudal anterior	95808	95961	89569	89722	< .001
Cingulate, rostral anterior	96104	96131	95818	95971	< .001
Posterior cingulate	108158	108317	108328	108480	< .001
Postcentral gyrus	102141	102293	102241	102394	< .001

Note: In all children (n=3198), unweighted models were compared to IPAW-weighted model models using the likelihood ratio test. A p-value < .05 suggests an improved model fit of the unweighted models. Models were adjusted for child sex and age at the neuroimaging assessment, maternal age at intake, maternal national origin, marital status, maternal education level, maternal substance use (tobacco, cannabis, alcohol), benzodiazepine usage during pregnancy, monthly household income and ICV for subcortical volumes. AIC = Akaike information criterion. BIC = Bayesian information criterion.

[a] Intracranial volume was additionally adjusted for.

eTable 6. Impact of Potentially Influential Observations on Primary Outcomes

Brain volumes	Altered Test Statistics[a]	Changed Significance
Cerebral white matter volume		
-Postnatal depression exposure only		
Observation 1	-1.67	NO
Amygdala		
-SSRI use before pregnancy		
Observation 1	1.51	NO
Observation 2	-1.76	NO
-Prenatal depression exposure		
Observation 1	1.04	NO
-Postnatal Depression only		
Observation 1	1.54	NO
Hippocampus		
-Prenatal SSRI exposure		
Observation 1	1.34	NO
-Postnatal depression exposure only		
Observation 1	1.61	NO

Note: To assess the potential impact of influential observations on the statistical significance of the primary outcomes, we conducted additional analyses, in which we tested the effect of deleting each of the potentially influential observations (as indicated by the red circles in eFigure 5) on the level of significance for each model. The results of these analyses revealed that none of the potential influential observations had a impacted the level of significance in any of the models. Thus, the results of our primary analyses are robust to the presence of influential observations.

[a] t-values were used as a test for significance at the commonly used threshold values of <-1.96 or >1.96, as specified in the eMethods section.

eReferences

1. Derogatis LR. Brief symptom inventory. *European Journal of Psychological Assessment*. Published online 1978.

2. Marroun HE, White TJH, Knaap NJF van der, et al. Prenatal exposure to selective serotonin reuptake inhibitors and social responsiveness symptoms of autism: population-based study of young children. *The British Journal of Psychiatry*. 2014;205(2):95-102. doi:10.1192/bjp.bp.113.127746

3. Andrews G, Peters L. The psychometric properties of the Composite International Diagnostic Interview. *Soc Psychiatry Psychiatr Epidemiol*. 1998;33(2):80-88. doi:10.1007/s001270050026

4. Brown MA, Lindheimer MD, de Swiet M, Assche AV, Moutquin JM. The Classification and Diagnosis of the Hypertensive Disorders of Pregnancy: Statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertension in Pregnancy*. 2001;20(1):ix-xiv. doi:10.3109/10641950109152635

5. White T, Marroun HE, Nijs I, et al. Pediatric population-based neuroimaging and the Generation R Study: the intersection of developmental neuroscience and epidemiology. *Eur J Epidemiol*. 2013;28(1):99-111. doi:10.1007/s10654-013-9768-0

6. White T, Muetzel RL, El Marroun H, et al. Paediatric population neuroimaging and the Generation R Study: the second wave. *Eur J Epidemiol*. 2018;33(1):99-125. doi:10.1007/s10654-017-0319-y

7. White T, Jansen PR, Muetzel RL, et al. Automated quality assessment of structural magnetic resonance images in children: Comparison with visual inspection and surface-based reconstruction. *Hum Brain Mapp*. 2018;39(3):1218-1231. doi:10.1002/hbm.23911

8. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, eds. *Selected Papers of Hirotugu Akaike*. Springer Series in Statistics. Springer; 1998:199-213. doi:10.1007/978-1-4612-1694-0_15

9. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*. 2004;11(1):192-196. doi:10.3758/BF03206482

10. Spilke J, Piepho HP, Hu X. Analysis of Unbalanced Data by Mixed Linear Models Using the mixed Procedure of the SAS System. *Journal of Agronomy and Crop Science*. 2005;191(1):47-54. doi:10.1111/j.1439-037X.2004.00120.x

11. Luke SG. Evaluating significance in linear mixed-effects models in R. *Behav Res*. 2017;49(4):1494-1502. doi:10.3758/s13428-016-0809-y

12. Cole SR, Hernán MA. Constructing Inverse Probability Weights for Marginal Structural Models. *American Journal of Epidemiology*. 2008;168(6):656-664. doi:10.1093/aje/kwn164

13. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ. Selection bias due to loss to follow up in cohort studies. *Epidemiology*. 2016;27(1):91-97. doi:10.1097/EDE.0000000000000409

14. Nieuwenhuis R, Grotenhuis M te, Pelzer B. influence.ME: Tools for Detecting Influential Data in Mixed Effects Models. *The R Journal*. 2012;4(2):38-47.

15. Team RC. R Core Team R: a language and environment for statistical computing. *Foundation for Statistical Computing*. Published online 2020.

16. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x