ANNALS of Clinical and Translational Neurology
Open Access
ANA American Neurological Association since 1875

RESEARCH ARTICLE

Caffeine for apnea of prematurity and brain development at 11 years of age

Claire E. Kelly[1,2], Wenn Lynn Ooi[1,2], Joseph Yuan-Mou Yang[2,3,4], Jian Chen[2], Chris Adamson[2], Katherine J. Lee[1,5,6], Jeanie L. Y. Cheong[1,7,8], Peter J. Anderson[1,9], Lex W. Doyle[1,6,7,8] & Deanne K. Thompson[1,2,6,10]

[1]Victorian Infant Brain Studies, Murdoch Children's Research Institute, Melbourne, Australia
[2]Developmental Imaging, Murdoch Children's Research Institute, Melbourne, Australia
[3]Department of Neurosurgery, The Royal Children's Hospital, Melbourne, Australia
[4]Neuroscience Research, Murdoch Children's Research Institute, Melbourne, Australia
[5]Clinical Epidemiology & Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia
[6]Department of Paediatrics, The University of Melbourne, Melbourne, Australia
[7]Department of Neonatal Services, The Royal Women's Hospital, Melbourne, Australia
[8]Department of Obstetrics and Gynaecology, The University of Melbourne, Melbourne, Australia
[9]Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Melbourne, Australia
[10]Florey Institute of Neuroscience and Mental Health, Melbourne, Australia

Correspondence
Claire Kelly, Victorian Infant Brain Studies (VIBeS), Murdoch Children's Research Institute, The Royal Children's Hospital, 50 Flemington Road, Parkville, Victoria, Australia, 3052. E-mail: claire.kelly@mcri.edu.au

Funding Information
This work was supported by the National Health and Medical Research Council of Australia (Project Grants No. 237117 and No. 108706; Program Grant No. 606789; Centre of Research Excellence No. 1060733; Fellowship No. 1081288 to PJA, No. 1127984 to KJL, No. 1141354 to JLYC and No. 1085754 to DKT). This work was also generously supported by The Royal Children's Hospital Foundation (Grant no. RCH1000) devoted to raising funds for research at The Royal Children's Hospital, as well as the Murdoch Children's Research Institute, the Royal Children's Hospital, Department of Paediatrics at The University of Melbourne and the Victorian Government's Operational Infrastructure Support Program

Received: 8 July 2018; Accepted: 10 July 2018

Annals of Clinical and Translational Neurology **2018; 5(9): 1112–1127**

doi: 10.1002/acn3.628

Abstract

Objective: Caffeine therapy for apnea of prematurity has been reported to improve brain white matter microstructure at term-equivalent age, but its long-term effects are unknown. This study aimed to investigate whether caffeine affects (1) brain structure at 11 years of age, and (2) brain development from term-equivalent age to 11 years of age, compared with placebo. **Methods**: Preterm infants born ≤1250 g were randomly allocated to caffeine or placebo. Magnetic resonance imaging (MRI) was performed on 70 participants (33 caffeine, 37 placebo) at term-equivalent age and 117 participants (63 caffeine, 54 placebo) at 11 years of age. Global and regional brain volumes and white matter microstructure were measured at both time points. **Results**: In general, there was little evidence for differences between treatment groups in brain volumes or white matter microstructure at age 11 years. There was, however, evidence that the caffeine group had a smaller corpus callosum than the placebo group. Volumetric brain development from term-equivalent to 11 years of age was generally similar between treatment groups. However, there was evidence that caffeine was associated with slower growth of the corpus callosum, and slower decreases in axial, radial, and mean diffusivities in the white matter, particularly at the level of the centrum semiovale, over time than placebo. **Interpretation**: This study suggests any benefits of neonatal caffeine therapy on brain structure in preterm infants weaken over time and are not clearly detectable by MRI at age 11 years, although caffeine may have long-term effects on corpus callosum development.

Introduction

Caffeine is used to treat cessation of breathing in infants born preterm (apnea of prematurity).[1] The Caffeine for Apnea of Prematurity (CAP) randomized controlled clinical trial established that caffeine has many benefits, and no harmful effects, on clinically important outcomes in infants born preterm and low birthweight.[1] Benefits included reduced rates of lung injury (bronchopulmonary dysplasia), retinopathy of prematurity and surgery for patent ductus arteriosus, and increased rates of survival free of major neurodevelopmental disabilities at 18–21 months of age.[1,2] The effects of caffeine were attenuated when the children were older, but caffeine was still associated with reduced rates of developmental coordination disorder at 5 years of age,[3] and improved motor outcomes, including manual dexterity and balance, at 5 and 11 years of age.[4,5]

The mechanism by which caffeine improves neurodevelopmental outcomes in preterm infants is not clear.[2] Several animal studies have reported that caffeine has beneficial effects on the brain, but many others warn of adverse effects.[6] In a subgroup of participants from the CAP trial, caffeine was associated with improved white matter (WM) microstructure in preterm infants at approximately term-equivalent age (38–44 weeks' gestational age), which may explain how caffeine improved early neurodevelopmental outcomes.[7] However, it is not known whether this improvement to brain structure persists or weakens in the long-term.

The primary aim of this study was to investigate the relationship between neonatal caffeine treatment and brain structure at 11 years of age in a subgroup of the CAP trial, including the structure of both global brain regions and specific brain regions involved in motor function, given the previously reported benefit of caffeine on long-term motor outcomes in the overall CAP trial.[5] The secondary aim was to investigate the relationship between neonatal caffeine treatment and longitudinal brain development from term-equivalent to 11 years of age.

Materials and Methods

Participants

The CAP trial enrolled 2006 infants with birthweights 500–1250 g, who were randomly allocated to caffeine or placebo.[1,2,4,5,7] Randomization was stratified according to study center. A subgroup ($n = 199$) of the infants enrolled in the CAP trial, who were cared for at the Royal Women's Hospital (RWH), Melbourne, is relevant to this study. Two of the 199 infants were excluded due to congenital abnormalities (craniosynostosis and Klinefelter syndrome). One hundred and seventy-two children (87% of the children recruited at RWH) were followed up at 11 years of age, along with an additional five children who were originally enrolled through other centers. Of these, 118 children had magnetic resonance imaging (MRI). One child was excluded from all subsequent analyses because of dental braces that compromised image quality. This left a maximum of 117 children (63 allocated to caffeine and 54 allocated to placebo) who were included in this study for the primary aim.

Of the 199 infants cared for at the RWH, 70 (33 caffeine, 37 placebo) had MRI at term-equivalent age, including structural MRI, with 28 (15 caffeine, 13 placebo) of those also having diffusion MRI, as previously reported.[7] All of the term-equivalent and 11-year MRI data were included in the secondary aim of the current study.

Ethics approval for recruitment and follow up was granted by the Human Research Ethics Committee at the RWH and written informed parental consent was necessary for participation at each follow up.

Data collection at age 11 years

Children underwent 3 Tesla MRI (Siemens Magnetom Trio-Tim syngo) at 11 years of age at the Royal Children's Hospital, Melbourne. T_1 images were acquired: repetition time (TR)/echo time (TE) 1950/2.24 msec, 0.9 mm isotropic voxels, field of view (FOV) 230 × 221 mm and flip angle 9 degrees. Diffusion images were acquired via two sequences: (1). TR/TE 7500/89 msec, FOV 240 × 240 mm, 2.5 mm isotropic voxels, b-value 1000 sec/mm^2, 25 diffusion-weighted gradient directions, five $b = 0$ sec/mm^2 volumes; (2). TR/TE 8000/112 msec, FOV 240 × 240 mm, 2.5 mm isotropic voxels, b-value = 3000 sec/mm^2, 45 diffusion-weighted gradient directions, six $b = 0$ sec/mm^2 volumes. All children had both of the diffusion sequences, but one child in the caffeine group did not complete the second ($b = 3000$ sec/mm^2) sequence.

Brain volumes

Volumes were generated from the T_1 images. The volumes of the intracranial cavity, total brain tissue and cerebrospinal fluid were calculated using Statistical Parametric Mapping (SPM; version 12).[8] The volume, surface area and thickness of the cortical gray matter (GM), WM volume and brainstem volume were calculated using FreeSurfer (version 5.3).[9–13] Subcortical and deep nuclear GM volumes [thalamus, basal ganglia nuclei, hippocampus and amygdala] were calculated using the Functional MRI of the Brain Software Library's (FSL's) Integrated Registration and Segmentation Tool (FIRST).[14] Cerebellum volume was calculated using SPM's Spatially Unbiased Infra-tentorial template (SUIT; version 3.0).[15,16] Area and thickness of the corpus callosum were calculated using

our previously described software pipeline,[17,18] which divided the corpus callosum into six subregions; genu, rostral body, anterior mid-body, posterior mid-body, isthmus, and splenium.[19] Imaging outputs were visually examined and manually edited as required.

Whole-brain white matter microstructure

Whole brain WM microstructure was analyzed using:

- FSL's (version 5.0.9) Tract-Based Spatial Statistics (TBSS).[20] This involved analyzing the $b = 1000$ sec/mm^2 sequence. Processing steps were: motion and eddy current distortion correction; diffusion tensor fitting to generate fractional anisotropy (FA), and axial (AD), radial (RD), and mean (MD) diffusivity images; alignment of all participants' images to a study-specific template; generation of a tract skeleton. Despite being commonly used, TBSS has methodological limitations, including that the metrics generated can be influenced by factors other than WM microstructure.[21] A complementary analysis using a more sophisticated WM modeling strategy was therefore carried out.
- MRtrix's (version 3) fixel-based analysis.[22] This involved analyzing the $b = 3000$ sec/mm^2 sequence. All recommended steps were followed. Fixel-based analysis enables investigation of the density and cross-section of WM fiber populations across the whole brain, thus providing metrics that are more specific to the WM microstructural environment compared with TBSS.[22]

White matter tractography

We reconstructed the corpus callosum, pyramidal sensorimotor, and cerebellar motor tracts (Fig. 1A). This involved drawing regions of interest on the diffusion images ($b = 3000$ sec/mm^2 sequence) and using a probabilistic fiber orientation distribution-based computer algorithm to reconstruct the tracts (Tournier et al. Improved probabilistic streamlines tractography by second order integration over fiber orientation distributions. *Proceedings of the International Society for Magnetic Resonance in Medicine* 2010;1670). Following tractography reconstructions, we calculated average FA, AD, RD, and MD of each of the tracts for every participant. When tracts were reconstructed using manually drawn regions of interest, we assessed intra-rater reliability using intra-class correlation coefficients, which were all ≥0.88, suggesting high intra-rater reliability.

Data collection for longitudinal analyses

At term-equivalent age, we acquired brain T_1-, T_2- and diffusion images during infants' sleep without sedation using a 1.5 Tesla MRI (General Electric Signa, GE Medical Systems) at the Royal Children's Hospital, Melbourne, as detailed previously.[7]

Based on the T_1 and T_2 images, we generated volumes of the intracranial cavity, cortical GM, WM, deep nuclear GM, cerebellum, brainstem, and hippocampus. We used software specifically designed for neonatal images,[23,24] or manual delineation for the hippocampus.[25] These data were used to analyze brain volumetric development.

Using the neonatal T_1 and diffusion images, we calculated the area of, and diffusion tensor values (FA, AD, RD, and MD) within, the corpus callosum and its six subregions, as previously described.[19] These data were used to analyze corpus callosum development.

In our previous study at term-equivalent age, we calculated diffusion tensor values from within manually drawn regions of interest.[7] Three regions (one in the anterior, one in the central and one in the posterior WM) were placed bilaterally on both an inferior brain slice at the level of the mid-thalamus, and a superior brain slice at the level of the centrum semiovale, that is, a total of 12 regions per participant.[7] To enable a longitudinal analysis, we drew the same regions on the 11-year diffusion images, for the subset of participants with diffusion images at both time points (Fig. 1B). This left us with diffusion values (FA, AD, RD, MD) from 12 regions per participant per time point. Intra-rater reliability for the manually drawn 11-year regions was high (all intra-class correlations ≥0.8, except FA within the right superior central region [0.71] and AD within the right inferior posterior region [0.77]).

Statistical analyses

Cross-sectional analysis at age 11 years (primary aim)

For the whole-brain analyses, we compared cortical GM surface area, thickness, and volume between treatment groups using FreeSurfer's Query, Design, Estimate and Contrast (QDEC).[26] We compared WM diffusion tensor values (derived from TBSS) between treatment groups using FSL's randomize, with 5000 permutations and threshold-free cluster enhancement.[27] We compared fiber density and cross-section (derived from the fixel-based analysis) between treatment groups using connectivity-based fixel enhancement.[28] These analyses were adjusted for age and sex, and corrected for multiple comparisons using the false discovery rate method (for QDEC) or the family-wise error rate method (for TBSS and fixel-based analyses).

Global and regional volume and WM microstructure measures were compared between treatment groups using

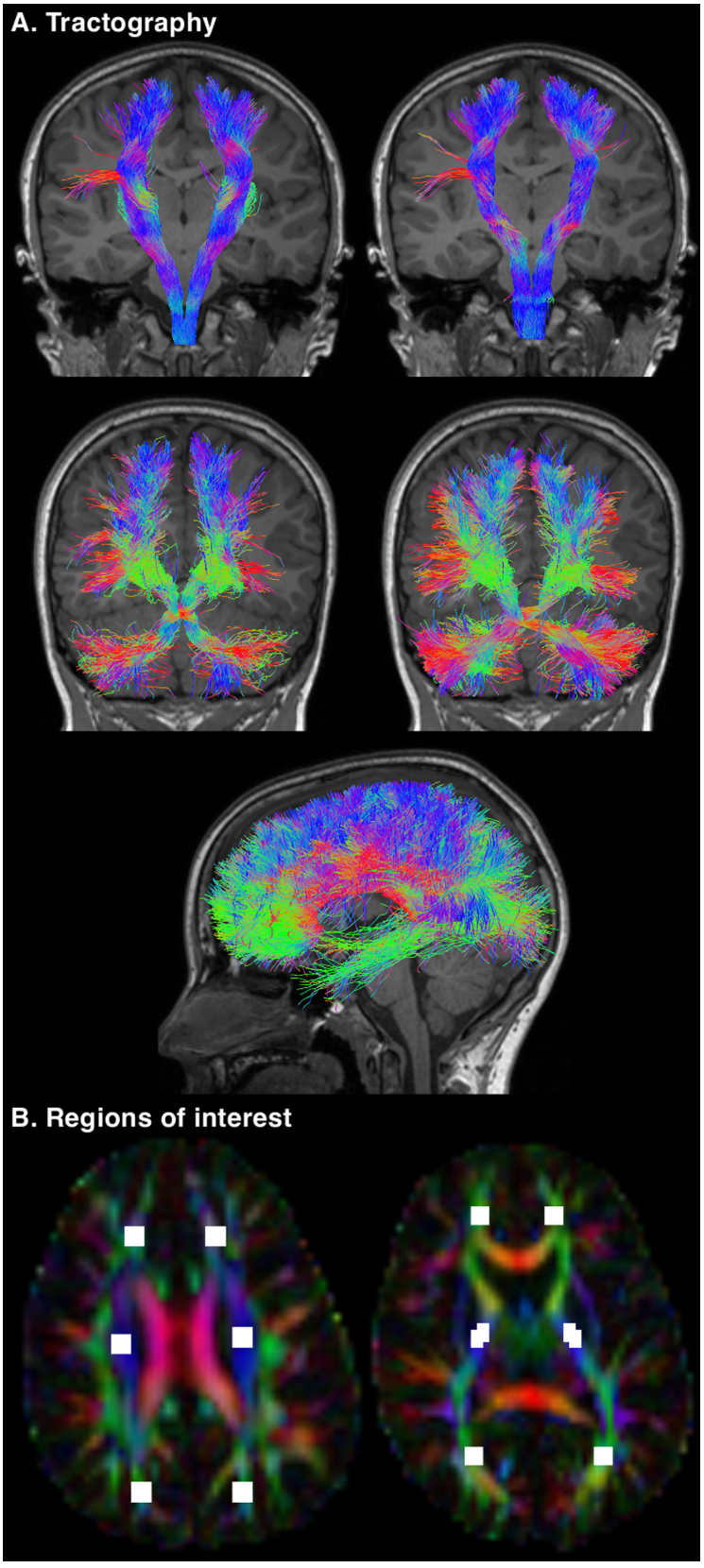


Figure 1. Illustration of some of the data generated for this study. (A) Tractography performed using the 11-year diffusion images. *Top row:* the pyramidal sensorimotor tracts. The corticospinal tract (left) was reconstructed using seed regions in the pons and inclusion regions in the posterior limb of the internal capsule. The somatosensory tract (right) was reconstructed using seed regions in the medial lemniscus, inclusion regions in the thalamus and the combined pericentral cortices (precentral, paracentral, and postcentral cortices), and exclusion regions in the corticospinal tract portion of the pons. *Middle row:* the cerebellar motor tracts. The cerebellar-thalamo-cortical tract (left) was reconstructed using seed regions in the dentate nucleus and inclusion regions in the decussation of the superior cerebellar peduncle, and the contralateral red nucleus, thalamus, and prefrontal cortex. The cortico-ponto-cerebellar tract (right) was reconstructed using seed regions in the middle cerebellar peduncle, inclusion regions in the contralateral cerebral peduncle, anterior limb of the internal capsule and prefrontal cortex, and exclusion regions in the decussation of the superior cerebellar peduncle and medial lemniscus. *Bottom row:* the corpus callosum tracts, reconstructed using automatically generated[17,18] seed regions in the corpus callosum, with the addition of exclusion regions at the brainstem, cerebral peduncle and thalamus. (B) Regions of interest drawn on the 11-year diffusion images. Six regions were placed on a single superior brain slice at the level of the centrum semiovale (left) and six regions were placed on a single inferior brain slice at the level of the mid-thalamus (right). These regions of interest were intended to reproduce the regions of interest drawn on the neonatal diffusion images, to enable a comparison of white matter microstructure between the time points. The neonatal regions of interest are shown in Figure 1 of our previous publication.[7]

separate linear regression models, with adjustment for age and sex (Stata version 14). Models were fitted using generalized estimating equations to account for correlations between data from multiple births, which represent a large proportion of our sample (Table 1). Some outcomes were measured in both brain hemispheres, in which case we included data from both hemispheres in a single model, and instead of allowing for correlations between data from multiple births, we allowed for correlations between data from individuals, which we expected to be more important. There was little evidence for any interactions between group and hemisphere, so the interaction terms were removed and results are reported combined across hemispheres.

Longitudinal analysis (secondary aim)

We compared the rate of change in the global and regional volume and WM microstructural measures from term-equivalent to 11 years of age between the caffeine and placebo groups using separate mixed effects models (Stata version 14). Models included a fixed effect of age and group, a random effect to allow for the repeated observations within individuals, and adjustment for sex. Age-by-group interactions were included. The diffusion values from regions of interest were measured in both brain hemispheres at both time points, and values from both hemispheres were analyzed using a single model that included a fixed effect of hemisphere. There was little evidence of any interactions between group, age and hemisphere, so these interaction terms were removed and results are reported combined across hemispheres.

For both aims, all the results were very similar before and after adjustment for total intracranial volume, so unadjusted results have been reported. Given the influence that bronchopulmonary dysplasia,[29] major neonatal brain injuries[30] and socioeconomic status[31] can have on brain development in preterm-born children, we also conducted three separate sensitivity analyses: (1) adjusting for bronchopulmonary dysplasia; (2) excluding the two children in our sample who had major neonatal brain injury (one had intraventricular hemorrhage (IVH) grade 3 and one had both IVH grade 3 and cystic periventricular leukomalacia (PVL)); (3) adjusting for maternal education level at birth.

Given the large number of comparisons performed, the analyses with global and regional volumes and WM microstructure measures were interpreted based on the overall patterns and magnitudes of findings, rather than specific *P*-values.

Results

Participants

Baseline characteristics were similar between participants in this study in the caffeine ($n = 63$) and placebo ($n = 54$) groups, except fewer participants in the caffeine group had bronchopulmonary dysplasia and treatment for a patent ductus arteriosus than in the placebo group (Table 1). This is consistent with the findings of the overall CAP trial.[1]

There was little evidence that cognitive and motor outcomes at 11 years of age differed between the caffeine ($n = 63$) and placebo ($n = 54$) participants in this study (Table 1). This is in contrast to the results of the overall CAP trial (total 920 participants), which found that the caffeine group had better motor outcomes at 11 years of age than the placebo group.[5] However, the magnitudes of the differences observed in motor scores in our subset (Table 1) were similar to those reported for the CAP trial overall,[5] suggesting that this difference in study results is most likely due to a lack of power in our substudy rather than systematic differences in the substudy population.

Of the seven children included in this study who had cerebral palsy at age 11 years (Table 1), none had cystic PVL or IVH grade 3 or 4 in the neonatal period.

Table 1. Summary of the perinatal and 11-year characteristics by treatment group.

Characteristic	Caffeine $n = 63$	Placebo $n = 54$	Mean difference (95% CI)	*P*-value
Around birth[1]				
Birth weight in grams	960 (188)	935 (184)	26 (−43, 94)	0.46
Gestational age at birth in weeks	27.7 (1.7)	27.4 (1.7)	0.3 (−0.3, 0.9)	0.36
Male sex, *n* (%)	28 (44.4)	30 (55.6)	OR 0.6, 95% CI 0.3, 1.3	0.23
Multiple births, *n* (%)	21 (33.3)	16 (29.6)	OR 1.2, 95% CI 0.5, 2.6	0.67
Cystic periventricular leukomalacia, *n* (%)	0 (0.0)	1 (1.9)	NA	NA
Intraventricular hemorrhage grade 3/4, *n* (%)	0 (0.0)	2 (3.7)	NA	NA
Bronchopulmonary dysplasia, *n* (%)	18 (28.6)	33 (61.1)	OR 0.3, 95% CI 0.1, 0.6	0.001
Patent ductus arteriosus, *n* (%)	12 (19.1)	22 (40.7)	OR 0.3, 95% CI 0.1, 0.8	0.01
Corrected postmenstrual age at MRI in weeks	40.7 (1.1)[3]	40.9 (1.4)[4]	−0.1 (−0.9, 0.6)	0.71
Mother completed post-secondary school education, *n* (%)	25 (41.0)[7]	18 (33.3)	OR 1.4, 95% CI 0.6, 3.0	0.40
At 11 years[2]				
Corrected age at MRI in years	11.3 (0.4)	11.3 (0.4)	0.03 (−0.1, 0.2)	0.69
Corrected age at neurodevelopmental assessment in years	11.3 (0.4)	11.3 (0.4)	0.03 (−0.1, 0.2)	0.69
IQ				
Full scale	95.2 (12.4)	96.9 (13.4)[5]	−1.6 (−6.4, 3.1)	0.50
Verbal comprehension index	97.3 (11.5)	97.5 (13.7)[5]	−0.2 (−4.9, 4.4)	0.93
Perceptual reasoning index	94.5 (15.5)	96.8 (15.7)	−2.3 (−8.0, 3.4)	0.43
Wide range achievement test				
Word reading standard score	98.5 (15.2)	97.1 (13.1)	1.5 (−3.8, 6.7)	0.58
Sentence comprehension standard score	100.0 (13.9)	100.2 (15.9)	−0.2 (−5.7, 5.2)	0.94
Spelling standard score	96.9 (13.4)	95.5 (14.9)	1.4 (−3.8, 6.6)	0.59
Math computation standard score	86.4 (14.1)	89.1 (15.4)	−2.8 (−8.2, 2.6)	0.31
Digit span				
Forwards	7.3 (2.7)[5]	7.8 (2.8)	−0.5 (−1.5, 0.6)	0.38
Backwards	8.4 (2.5)[5]	8.4 (2.2)	0.1 (−0.8, 0.9)	0.88
Beery VMI				
Visual perception standard score	95.9 (11.6)	94.1 (13.9)	1.8 (−2.9, 6.4)	0.45
Motor coordination standard score	92.1 (12.6)	90.4 (15.5)	1.7 (−3.4, 6.8)	0.51
Movement ABC				
Total standard score	9.2 (3.2)[6]	8.7 (2.9)[5]	0.5 (−0.6, 1.7)	0.37
Manual dexterity standard score	8.1 (3.1)[7]	8.1 (2.9)[5]	0.03 (−1.1, 1.2)	0.95
Aiming/catching standard score	10.3 (3.8)[5]	9.6 (0.5)[5]	0.7 (−0.7, 2.0)	0.33
Balance standard score	10.1 (3.4)[7]	9.6 (2.6)[5]	0.6 (−0.6, 1.7)	0.33
Cerebral palsy, *n* (%)	5 (7.9)	2 (3.7)	OR 2.2, 95% CI 0.4, 12.1	0.35
Blindness, *n* (%)	0 (0.0)	0 (0.0)	NA	NA
Deafness, *n* (%)	1 (4.8)[3]	2 (9.5)[3]	OR 0.5, 95% CI 0.04, 5.7	0.56

Data are mean (SD), unless otherwise stated.

ABC, Assessment battery for children; CI, confidence interval; OR, odds ratio; MRI, magnetic resonance imaging; NA, not applicable; VMI, visual motor integration test.

[1]Perinatal data were collected by chart review. Intraventricular hemorrhage was defined as described previously.[40] Bronchopulmonary dysplasia was defined as requirement for supplemental oxygen at 36 weeks' postmenstrual age. Patent ductus arteriosus was defined as any case requiring medical or surgical treatment. We did not have data on retinopathy of prematurity.

[2]Data on functional outcomes at 11 years of age (cognitive, academic, motor and behavioral outcomes) were collected as detailed previously.[5] We did not collect any data on pubertal development.

[3]$n = 21$.

[4]$n = 22$.

[5]$n = 1$ with missing data.

[6]$n = 3$ with missing data.

[7]$n = 2$ with missing data.

Radiologists reviewed the MRI of these seven children at age 11 years, and reported that none of these children had features on their MRI that warranted clinical follow up.

Perinatal characteristics were similar between the $n = 117$ participants included in this study and the remaining $n = 85$ participants from the RWH who were excluded from this study (Table S1). Perinatal characteristics of our subset ($n = 117$) were also similar to those previously reported for the overall CAP cohort ($n = 920$).[5]

Cross-sectional analysis at age 11 years

Whole brain

Using FreeSurfer's QDEC, there was little evidence that cortical GM area, thickness and volume differed between treatment groups. Using TBSS, there was little evidence that FA, AD, RD, and MD differed between treatment groups (that is, there were no vertices for QDEC and no voxels for TBSS that differed between groups at $P < 0.05$, multiple comparison corrected, hence no data can be shown). Using the fixel-based analysis, there was evidence that fiber cross-section was lower in the caffeine group compared with the placebo group in 450 fixels (0.2% of the total 210998 fixels), located predominantly in the

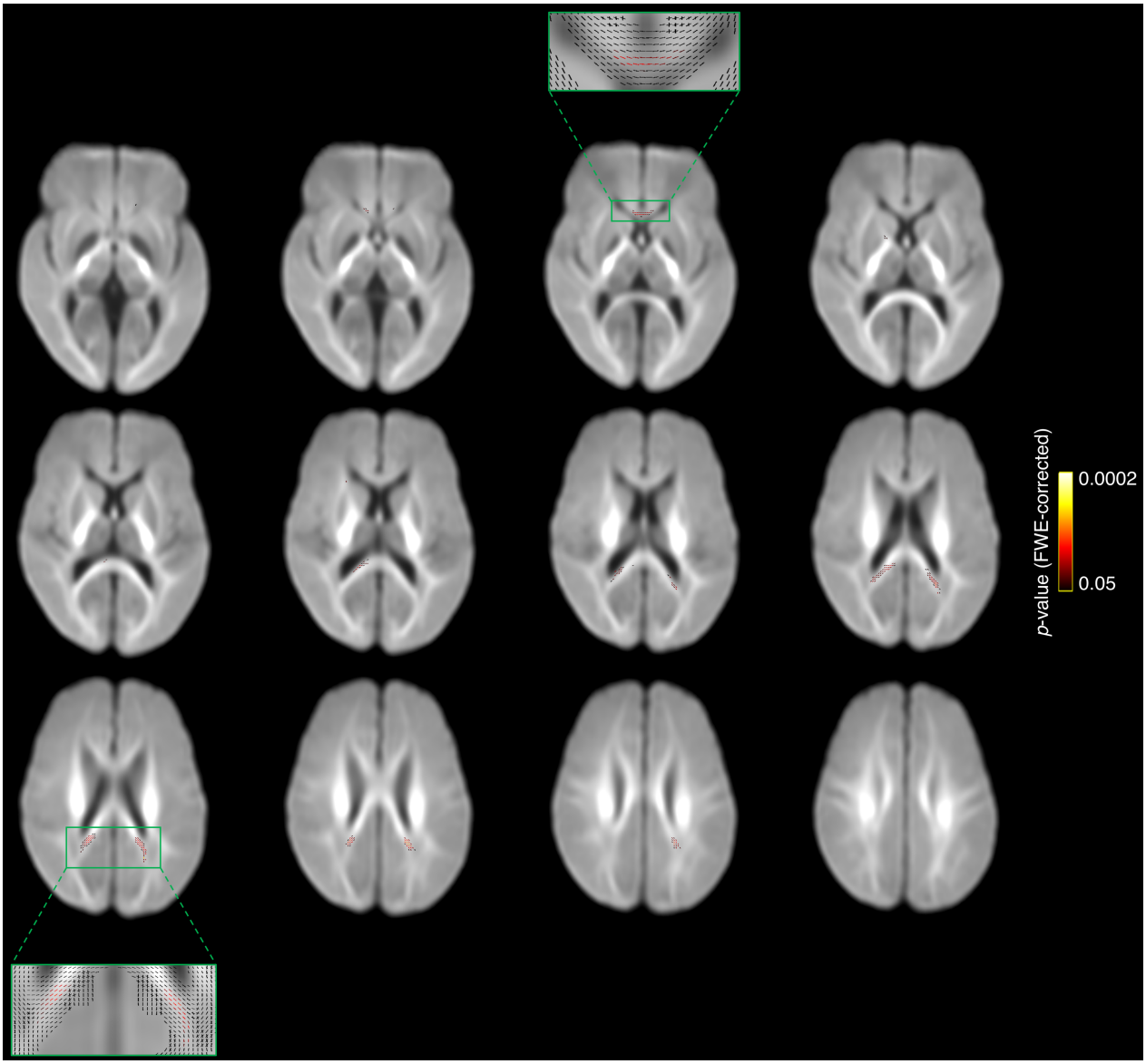


Figure 2. Results of the whole brain white matter microstructure analysis (fixel-based analysis) at 11 years of age. The fiber populations that had a lower fiber cross-section in the caffeine group compared with the placebo group at $P < 0.05$, family-wise error rate (FWE) corrected, are shown in red-yellow and overlaid on the study-specific fiber-orientation distribution template. These fibers were located mostly in the isthmus or splenium of the corpus callosum, as well as the genu of the corpus callosum, and a very small number were in the left anterior limb of the internal capsule. We have presented every second axial slice (in the range of the significant results). We have zoomed in on representative slices to show the fiber orientation distribution in these slices, which shows that the significant results were located in single fiber regions. These results are based on an analysis with $n = 116$ children (62 caffeine, 54 placebo), that is, one less than the total $n = 117$. One child was excluded from this analysis as they did not complete the $b = 3000$ sec/mm^2 diffusion sequence.

corpus callosum isthmus or splenium subregions, and to a lesser extent the genu, and a very small number were located in the right anterior limb of the internal capsule (Fig. 2). Only 6 fixels (0.0003% of the total fixels) in the left anterior limb of the internal capsule had lower values of the combined measure of fiber density and cross-section in the caffeine group compared with the placebo group (data not shown). There was little evidence that fiber density differed between treatment groups. Results were very similar in the three sensitivity analyses, but the number of fixels increased slightly in each sensitivity analysis compared with the original analysis (Table S2).

Global and regional volumes and white matter microstructure

Overall, there was little evidence that global brain volumes (Fig. 3A), cortical, subcortical and deep nuclear GM volumes and cerebellar volumes (Fig. 3B), and microstructure and volume of primary sensorimotor and cerebellar tracts (Fig. 3C) differed between treatment groups. However, there was some evidence that volume of the cerebellar tracts was lower in the caffeine group compared with the placebo group (Fig. 3C). All these results were very similar in the three sensitivity analyses (data not shown).

There was some evidence that the corpus callosum was thinner along much of its length in the caffeine group compared with the placebo group (Fig. 4A), and this difference was strongest in the rostral body and isthmus subregions (Fig. 4A and B). There was also evidence that corpus callosum area was smaller in the caffeine group compared with the placebo group, particularly in the genu, rostral body, and isthmus (Fig. 4B). Microstructure and volume of corpus callosum tracts were generally similar between treatment groups, although there was some evidence that RD and MD in all subregions, and particularly the genu, were higher in the caffeine group compared with the placebo group (Fig. 4B). These differences in the corpus callosum were generally very similar in the three sensitivity analyses; if anything, the strength and magnitude of the differences generally increased slightly after adjusting for bronchopulmonary dysplasia (data not shown).

Longitudinal analysis

There was little evidence that the rate of change in brain volumes varied between the caffeine and placebo groups (Fig. 5).

There was some evidence that the area of the corpus callosum increased more slowly in the caffeine group compared with the placebo group, particularly in the genu, rostral body, and isthmus (Fig. 6). There was also evidence that AD, RD, and MD in the corpus callosum tracts, particularly the genu, rostral body, and splenium, decreased more slowly in the caffeine group compared with the placebo group.

There was evidence that AD, RD, and MD (but not FA) within the regions of interest decreased more slowly in the caffeine group compared with the placebo group, and the evidence was stronger for the superior regions than the inferior regions (Fig. 7).

The results of the longitudinal analyses were very similar in the three sensitivity analyses (data not shown).

Discussion

Short-term neonatal caffeine treatment had little effect on the volume or microstructure of most of the brain 11 years later, except for the corpus callosum, which was smaller in the caffeine group. Additionally, caffeine was weakly associated with brain volumetric development from term-equivalent to 11 years of age, but caffeine was associated with slower growth of the corpus callosum over time, and slower decreases in WM diffusivities over time.

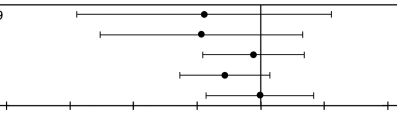
In our previous study, caffeine had little effect on brain volumes, but benefitted WM microstructure by reducing AD, RD, and MD in superior WM regions at the level of the centrum semiovale compared with placebo, at term-equivalent age.[7] In the current study, we showed that WM diffusivities decreased more slowly from term-equivalent to 11 years of age in the caffeine group compared with the placebo group, such that they generally did not differ between treatment groups at age 11 years. If we accept from previous studies that decreases in diffusivities, particularly RD and MD, are expected and beneficial during childhood development,[32,33] the slower

Figure 3. Global and regional brain volumes and white matter microstructure, contrasted between treatment groups at 11 years of age. Mean differences and *P*-values are from separate linear regression models for each MRI outcome, adjusted for age and sex of the child. The plots are visual representations of the mean differences and 95% confidence intervals (CI). For bilateral outcomes, data from the left and right hemispheres were analyzed in a single regression model, and results are presented as a single estimate for the difference for the two hemispheres, as there was little evidence of group by hemisphere interactions. Units are cm^3 for volumes and $\times 10^{-3}$ mm^2/sec for diffusivities. SD = standard deviation. The analyses are based on the total $n = 117$ (63 caffeine, 54 placebo) children, although some children are missing from some analyses because their images had artifact which affected that particular analysis: [a]$n = 3$ missing; [b]$n = 2$ missing; [c]$n = 1$ missing; [d]$n = 5$ missing; [e]$n = 4$ missing. The total numbers are lower for the tractography analyses ($n = 116$; 62 caffeine, 54 placebo) because one child in the caffeine group did not have $b = 3000$ sec/mm^2 diffusion images acquired and hence had to be excluded from the tractography analyses.

A. Global brain volumes

Brain region/tissue type	Caffeine, *n*=63 Mean	SD	Placebo, *n*=54 Mean	SD	Mean difference	95% CI Lower	95% CI Upper	*p*-value
Intracranial volume	1399	131	1429	109	-17.8	-57.9	22.2	0.38
Total brain tissue volume	1194	105	1222	89	-18.7	-50.5	13.2	0.25
Total cortical gray matter volume	545[a]	50	55[b]	43	-2.3	-18.2	13.7	0.78
Total white matter volume	399[a]	46	415[b]	44	-11.3	-25.5	2.9	0.12
Total cerebrospinal fluid volume	205	44	207	53	-0.3	-17.2	16.7	0.97

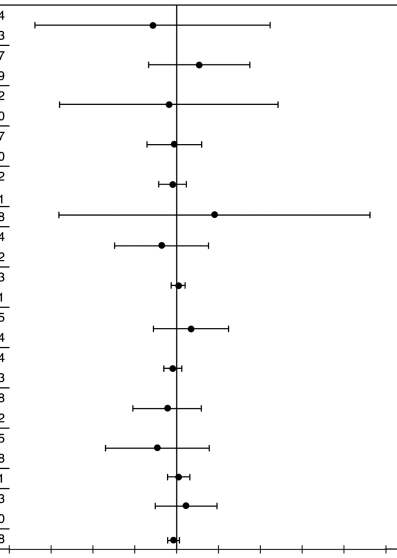
Plots of mean differences and 95% CI



B. Regional brain volumes

Brain region/tissue type	Sub-region	Caffeine, *n*=63 Mean	SD	Placebo, *n*=54 Mean	SD	Mean difference	95% CI Lower	95% CI Upper	*p*-value
Cortical gray matter (primary motor)	Precentral- left	14.88[a]	1.69	15.06[b]	1.84	-0.12	-0.68	0.45	0.69
	Precentral- right	14.52[a]	1.77	14.81[b]	1.73				
	Paracentral- left	4.44[a]	0.75	4.31[b]	0.77	0.11	-0.13	0.35	0.38
	Paracentral- right	4.87[a]	0.79	4.84[b]	0.69				
	Postcentral- left	11.03[a]	1.62	10.93[b]	1.42	-0.04	-0.56	0.48	0.89
	Postcentral- right	10.18[a]	1.59	10.52[b]	1.60				
Subcortical gray matter	Hippocampus- left	3.42	0.38	3.44	0.47	-0.01	-0.14	0.12	0.87
	Hippocampus- right	3.40	0.38	3.43	0.40				
	Amygdala- left	1.34	0.19	1.37	0.22	-0.02	-0.09	0.05	0.55
	Amygdala- right	1.27	0.20	1.28	0.21				
	Brainstem	19.12[a]	2.11	19.05[b]	2.28	0.18	-0.56	0.92	0.63
Deep nuclear gray matter	Thalamus- left	7.73	0.61	7.87	0.74	-0.07	-0.30	0.15	0.53
	Thalamus- right	7.54	0.59	7.61	0.72				
	Accumbens- left	0.49	0.10	0.48	0.13	0.01	-0.03	0.04	0.69
	Accumbens- right	0.43	0.07	0.43	0.11				
	Caudate- left	3.62	0.48	3.54	0.55	0.07	-0.11	0.25	0.46
	Caudate- right	3.64	0.53	3.61	0.54				
	Pallidum- left	1.59	0.14	1.62	0.14	-0.02	-0.06	0.02	0.39
	Pallidum- right	1.60	0.13	1.62	0.13				
	Putamen- left	4.80	0.48	4.85	0.58	-0.05	-0.21	0.12	0.58
	Putamen- right	4.71	0.46	4.82	0.52				
Cerebellar gray matter	Hemispheres- left	58.88	0.73	58.92	0.55	-0.09	-0.34	0.16	0.47
	Hemispheres- right	60.46	0.93	60.56	0.58				
	Vermis	6.07	0.15	6.07	0.11	0.01	-0.04	0.06	0.72
Cerebellar white matter	Hemispheres- left	4.10	0.44	4.08	0.33	0.04	-0.10	0.19	0.55
	Hemispheres- right	3.84	0.50	3.80	0.40				
	Vermis	0.25	0.09	0.27	0.08	-0.01	-0.04	0.01	0.30

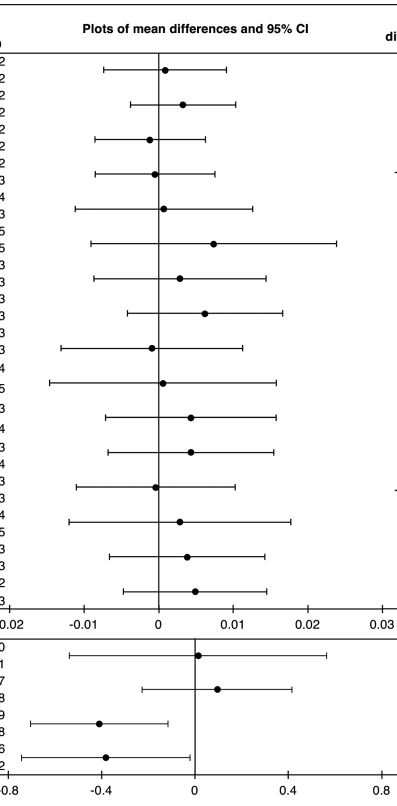
Plots of mean differences and 95% CI



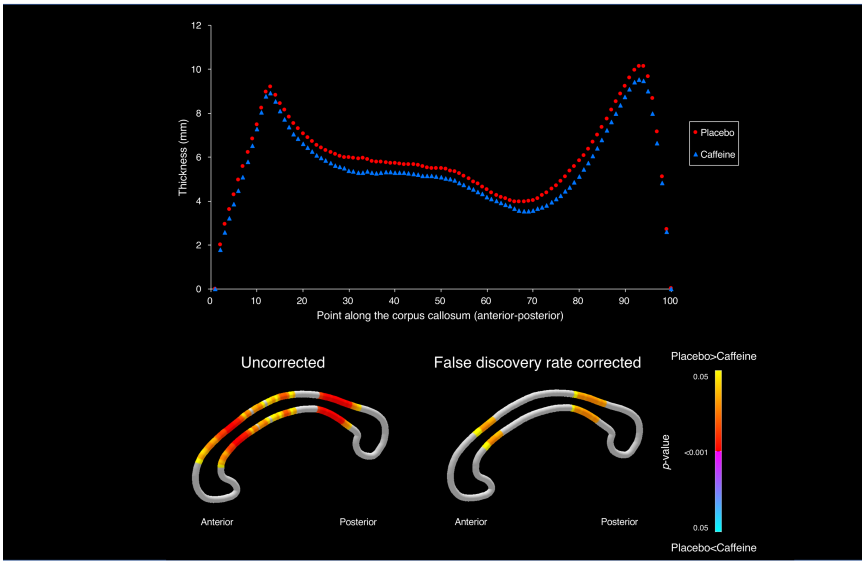
C. White matter tracts

Tract measure	Tract	Caffeine, *n*=62 Mean	SD	Placebo, *n*=54 Mean	SD	Mean difference	95% CI Lower	95% CI Upper	*p*-value
Fractional anisotropy	Corticospinal- left	0.46	0.03	0.45	0.02	0.001	-0.007	0.009	0.84
	Corticospinal- right	0.45	0.03	0.45	0.02				
	Somatosensory- left	0.42[c]	0.02	0.42	0.02	0.003	-0.004	0.010	0.36
	Somatosensory- right	0.42[c]	0.02	0.41	0.02				
	Cerebellar-thalamo-cortical - left	0.40[a]	0.02	0.40[c]	0.02	-0.001	-0.009	0.006	0.76
	Cerebellar-thalamo-cortical - right	0.40[a]	0.03	0.41[c]	0.02				
	Cortico-ponto-cerebellar- left	0.37[d]	0.02	0.38[e]	0.02	-0.0005	-0.009	0.008	0.90
	Cortico-ponto-cerebellar- right	0.37[d]	0.03	0.37[e]	0.03				
Axial diffusivity	Corticospinal- left	1.25	0.03	1.25	0.04	0.001	-0.011	0.013	0.91
	Corticospinal- right	1.26	0.04	1.26	0.03				
	Somatosensory- left	1.27[c]	0.05	1.26	0.05	0.007	-0.009	0.024	0.38
	Somatosensory- right	1.28[c]	0.05	1.27	0.05				
	Cerebellar-thalamo-cortical - left	1.20[a]	0.03	1.19[c]	0.03	0.003	-0.009	0.014	0.63
	Cerebellar-thalamo-cortical - right	1.19[a]	0.04	1.19[c]	0.03				
	Cortico-ponto-cerebellar- left	1.14[d]	0.03	1.14[e]	0.03	0.006	-0.004	0.017	0.24
	Cortico-ponto-cerebellar- right	1.15[d]	0.03	1.14[e]	0.03				
Radial diffusivity	Corticospinal- left	0.60	0.04	0.60	0.03	-0.001	-0.013	0.011	0.88
	Corticospinal- right	0.61	0.04	0.61	0.03				
	Somatosensory- left	0.66[c]	0.04	0.66	0.04	0.001	-0.015	0.016	0.94
	Somatosensory- right	0.66[c]	0.04	0.66	0.05				
	Cerebellar-thalamo-cortical - left	0.64[a]	0.03	0.63[c]	0.03	0.004	-0.007	0.016	0.46
	Cerebellar-thalamo-cortical - right	0.63[a]	0.04	0.63[c]	0.04				
	Cortico-ponto-cerebellar- left	0.64[d]	0.03	0.63[e]	0.03	0.004	-0.007	0.015	0.45
	Cortico-ponto-cerebellar- right	0.65[d]	0.04	0.64[e]	0.04				
Mean diffusivity	Corticospinal- left	0.81	0.03	0.82	0.03	-0.0004	-0.011	0.010	0.94
	Corticospinal- right	0.83	0.03	0.82	0.03				
	Somatosensory- left	0.86[c]	0.04	0.86	0.04	0.003	-0.012	0.018	0.71
	Somatosensory- right	0.87[c]	0.04	0.87	0.05				
	Cerebellar-thalamo-cortical - left	0.82[a]	0.03	0.82[c]	0.03	0.004	-0.007	0.014	0.47
	Cerebellar-thalamo-cortical - right	0.82[a]	0.03	0.81[c]	0.03				
	Cortico-ponto-cerebellar- left	0.81[d]	0.03	0.80[e]	0.02	0.005	-0.005	0.014	0.32
	Cortico-ponto-cerebellar- right	0.81[d]	0.03	0.81[e]	0.03				
Volume	Corticospinal- left	16.92	1.79	17.02	1.60	0.01	-0.54	0.56	0.97
	Corticospinal- right	17.28	1.72	17.24	1.61				
	Somatosensory- left	11.94[c]	0.92	11.72	0.97	0.09	-0.23	0.41	0.57
	Somatosensory- right	11.69[c]	1.06	11.80	1.18				
	Cerebellar-thalamo-cortical - left	13.16[a]	1.00	13.57[c]	0.99	-0.41	-0.70	-0.12	0.01
	Cerebellar-thalamo-cortical - right	12.85[a]	1.01	13.34[c]	1.08				
	Cortico-ponto-cerebellar- left	14.64[d]	1.12	15.15[e]	1.06	-0.38	-0.74	-0.02	0.04
	Cortico-ponto-cerebellar- right	14.43[d]	1.40	14.70[e]	1.22				

Plots of mean differences and 95% CI



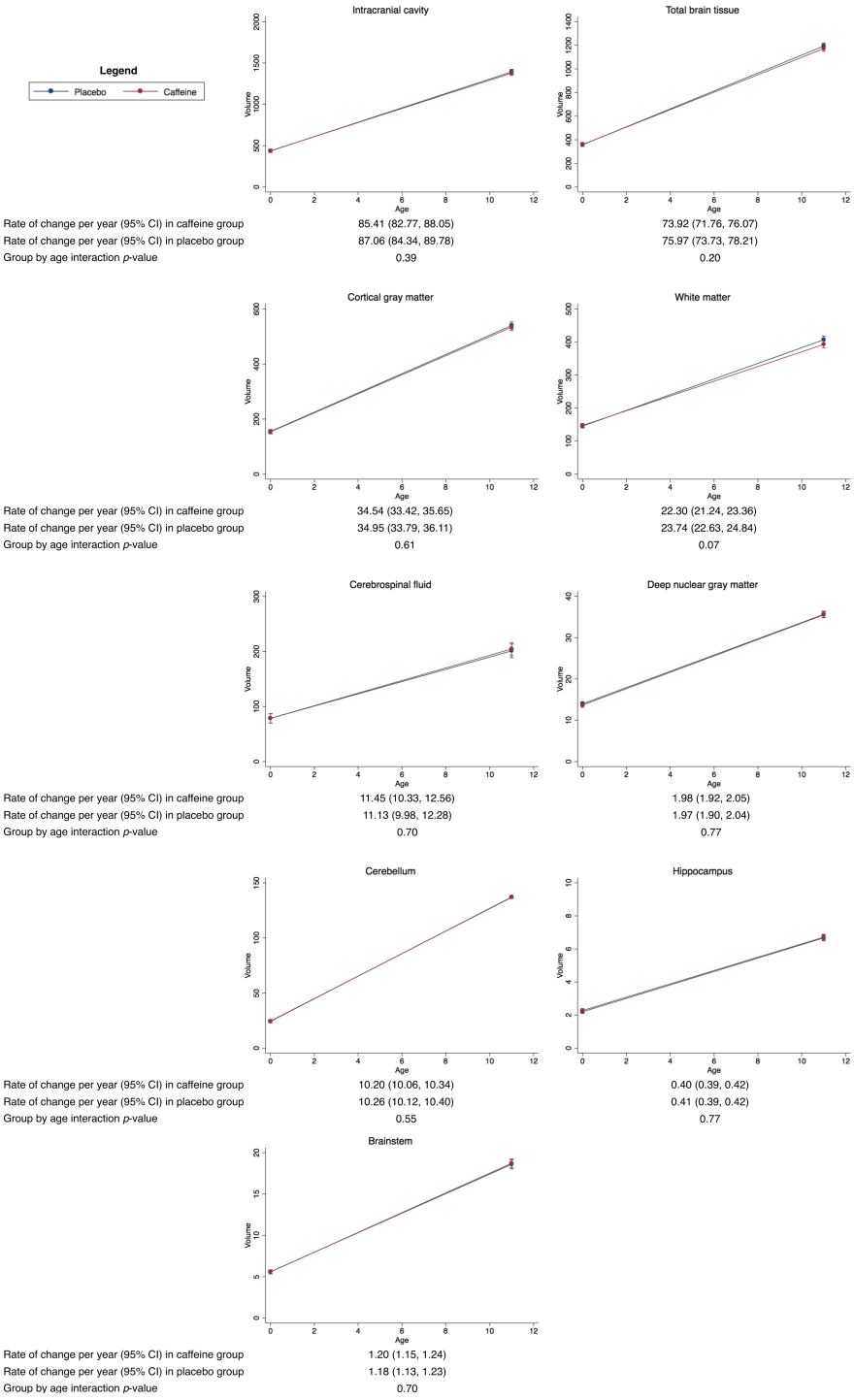
A. Corpus callosum thickness profile



B. Structure of the total corpus callosum and sub-regions

Corpus callosum measure	Sub-region	Caffeine, *n*=62 Mean	Caffeine, *n*=62 SD	Placebo, *n*=54 Mean	Placebo, *n*=54 SD	Plots of mean differences and 95% CI	Mean difference	95% CI lower	95% CI upper	*p*-value
Area	Total	4.91	1.03	5.43	0.84		-0.486	-0.812	-0.161	0.003
	Genu	1.35	0.31	1.52	0.28		-0.156	-0.261	-0.052	0.004
	Rostral body	0.69	0.16	0.77	0.14		-0.068	-0.118	-0.018	0.007
	Anterior mid body	0.60	0.14	0.65	0.11		-0.041	-0.085	0.003	0.067
	Posterior mid body	0.50	0.11	0.55	0.11		-0.042	-0.081	-0.003	0.033
	Isthmus	0.41	0.12	0.48	0.12		-0.065	-0.106	-0.023	0.002
	Splenium	1.34	0.33	1.46	0.31		-0.113	-0.231	0.005	0.061
Thickness	Total	0.51	0.08	0.55	0.08		-0.044	-0.072	-0.016	0.002
	Genu	0.49	0.08	0.53	0.07		-0.035	-0.060	-0.009	0.007
	Rostral body	0.55	0.10	0.61	0.11		-0.051	-0.085	-0.018	0.002
	Anterior mid body	0.52	0.10	0.55	0.09		-0.030	-0.064	0.005	0.090
	Posterior mid body	0.42	0.08	0.45	0.09		-0.028	-0.059	0.002	0.069
	Isthmus	0.41	0.11	0.47	0.12		-0.067	-0.107	-0.026	0.001
	Splenium	0.65	0.13	0.71	0.13		-0.054	-0.101	-0.008	0.023
FA	Total	0.49[a]	0.03	0.50[b]	0.02		-0.011	-0.022	-0.0003	0.043
	Genu	0.52[a]	0.05	0.54[b]	0.04		-0.017	-0.038	0.003	0.092
	Rostral body	0.44[a]	0.04	0.45[b]	0.04		-0.011	-0.026	0.004	0.143
	Anterior mid body	0.45[a]	0.04	0.47[b]	0.03		-0.010	-0.023	0.003	0.132
	Posterior mid body	0.47[a]	0.03	0.48[b]	0.02		-0.005	-0.016	0.006	0.377
	Isthmus	0.45[a]	0.04	0.46[b]	0.03		-0.007	-0.020	0.006	0.270
	Splenium	0.58[a]	0.06	0.60[b]	0.04		-0.018	-0.036	0.001	0.060
AD	Total	1.48[a]	0.07	1.48[b]	0.05		0.004	-0.020	0.028	0.767
	Genu	1.49[a]	0.09	1.49[b]	0.07		0.006	-0.023	0.036	0.675
	Rostral body	1.36[a]	0.08	1.37[b]	0.08		-0.005	-0.036	0.026	0.757
	Anterior mid body	1.37[a]	0.06	1.39[b]	0.07		-0.017	-0.043	0.009	0.202
	Posterior mid body	1.43[a]	0.07	1.42[a]	0.06		0.012	-0.014	0.037	0.372
	Isthmus	1.49[a]	0.10	1.49[b]	0.09		0.004	-0.032	0.041	0.823
	Splenium	1.73[a]	0.15	1.71[b]	0.12		0.019	-0.030	0.068	0.444
RD	Total	0.66[a]	0.06	0.64[b]	0.05		0.020	-0.002	0.042	0.070
	Genu	0.61[a]	0.07	0.58[b]	0.05		0.026	0.004	0.048	0.022
	Rostral body	0.66[a]	0.07	0.64[b]	0.06		0.015	-0.008	0.039	0.210
	Anterior mid body	0.65[a]	0.06	0.64[b]	0.06		0.010	-0.012	0.031	0.386
	Posterior mid body	0.67[a]	0.07	0.65[b]	0.06		0.016	-0.008	0.040	0.179
	Isthmus	0.73[a]	0.09	0.72[b]	0.08		0.019	-0.013	0.052	0.240
	Splenium	0.65[a]	0.14	0.62[b]	0.10		0.035	-0.012	0.082	0.146
MD	Total	0.93[a]	0.06	0.92[b]	0.05		0.015	-0.007	0.036	0.176
	Genu	0.90[a]	0.06	0.88[b]	0.04		0.023	0.004	0.043	0.018
	Rostral body	0.89[a]	0.07	0.88[b]	0.06		0.008	-0.016	0.031	0.531
	Anterior mid body	0.89[a]	0.05	0.89[b]	0.06		0.0002	-0.021	0.021	0.983
	Posterior mid body	0.92[a]	0.07	0.91[b]	0.06		0.015	-0.008	0.038	0.212
	Isthmus	0.99[a]	0.09	0.97[b]	0.07		0.014	-0.017	0.046	0.373
	Splenium	1.01[a]	0.14	0.98[b]	0.10		0.030	-0.016	0.075	0.199
Tract volume	Total	41.68[a]	3.46	42.03[b]	3.69		-0.274	-1.588	1.039	0.682
	Genu	6.25[a]	0.81	6.10[b]	1.03		0.150	-0.187	0.487	0.379
	Rostral body	6.16[a]	0.76	6.31[b]	0.67		-0.142	-0.410	0.126	0.300
	Anterior mid body	6.40[a]	0.71	6.42[b]	0.75		-0.019	-0.290	0.253	0.894
	Posterior mid body	7.20[a]	0.93	7.43[b]	0.84		-0.213	-0.558	0.133	0.228
	Isthmus	7.71[a]	1.01	7.64[b]	1.03		0.066	-0.315	0.447	0.736
	Splenium	7.94[a]	0.82	8.13[b]	0.88		-0.166	-0.486	0.154	0.310

Figure 4. Structure of the corpus callosum, contrasted between treatment groups at 11 years of age, based on the corpus callosum segmentation. (A) Top: Mean thickness at each of 100 points along the corpus callosum; bottom: results of *t*-tests comparing thickness at each point between treatment groups, with colors indicating *P*-values. (B) Data from the total corpus callosum and six sub-regions. Mean differences are from separate linear regression models for each outcome, adjusted for age and sex of the child. The plots are visual representations of the mean differences and 95% confidence intervals (CI). Units are cm^2 for area, cm for thickness and $\times 10^{-3}$ mm^2/sec for diffusivities. SD = standard deviation. The analyses are based on a total n = 116 (62 caffeine, 54 placebo) children, that is, one less than the total n = 117 children. One child from the caffeine group was excluded from all the corpus callosum analyses because they had artefact on their structural image and because they did not have b = 3000 sec/mm^2 diffusion images acquired. Additional children were excluded from the corpus callosum tractography analysis, because their diffusion images had artefact or structural abnormalities which affected that particular analysis: [a]n = 2 missing; [b]n = 5 missing.



Rate of change per year (95% CI) in caffeine group	85.41 (82.77, 88.05)	73.92 (71.76, 76.07)
Rate of change per year (95% CI) in placebo group	87.06 (84.34, 89.78)	75.97 (73.73, 78.21)
Group by age interaction *p*-value	0.39	0.20
Rate of change per year (95% CI) in caffeine group	34.54 (33.42, 35.65)	22.30 (21.24, 23.36)
Rate of change per year (95% CI) in placebo group	34.95 (33.79, 36.11)	23.74 (22.63, 24.84)
Group by age interaction *p*-value	0.61	0.07
Rate of change per year (95% CI) in caffeine group	11.45 (10.33, 12.56)	1.98 (1.92, 2.05)
Rate of change per year (95% CI) in placebo group	11.13 (9.98, 12.28)	1.97 (1.90, 2.04)
Group by age interaction *p*-value	0.70	0.77
Rate of change per year (95% CI) in caffeine group	10.20 (10.06, 10.34)	0.40 (0.39, 0.42)
Rate of change per year (95% CI) in placebo group	10.26 (10.12, 10.40)	0.41 (0.39, 0.42)
Group by age interaction *p*-value	0.55	0.77
Rate of change per year (95% CI) in caffeine group	1.20 (1.15, 1.24)	
Rate of change per year (95% CI) in placebo group	1.18 (1.13, 1.23)	
Group by age interaction *p*-value	0.70	

Figure 5. Rates of change in brain volumes from term-equivalent age to 11 years of age, contrasted between treatment groups. Results are from separate mixed effects models for each brain region, adjusted for sex of the child. The plots show the estimated means and 95% confidence intervals (CI) per group for each time point and the rate of change between time points from the mixed models. The y-axes show the volumes and the x-axes show age. Units are cm^3 for volumes and years for age. The results are based on the total number of children who had usable MRI data at either term-equivalent age or 11 years of age; for the intracranial cavity, total brain tissue, cerebrospinal fluid, deep nuclear gray matter and cerebellum, n = 187 (n = 70 with term-equivalent data plus n = 117 with 11-year data); for the cortical gray matter, white matter and brainstem, n = 182 (n = 70 with term-equivalent data plus n = 112 with 11-year data); for the hippocampus, n = 179 (n = 62 with term-equivalent data plus n = 117 with 11-year data).

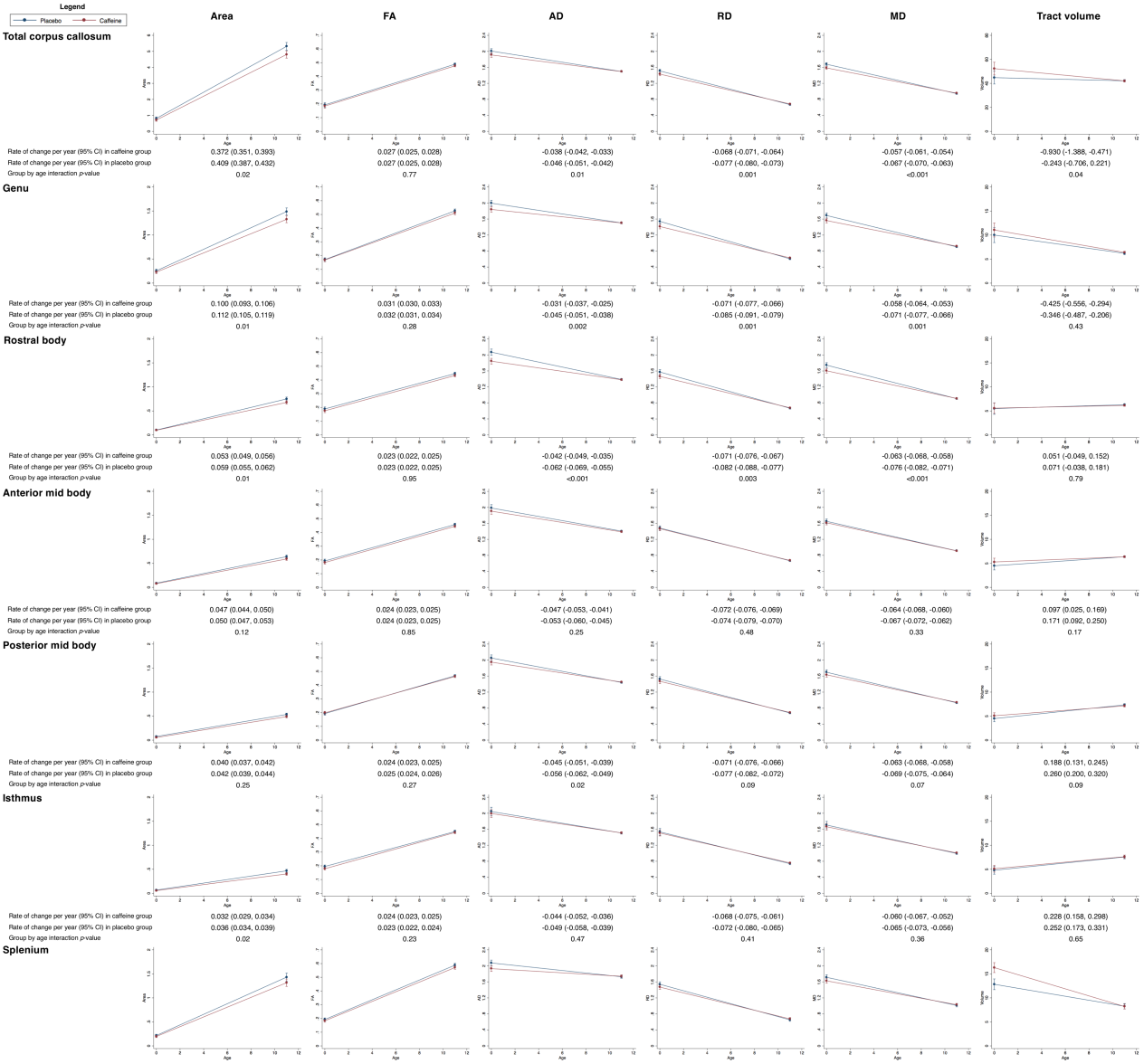


Figure 6. Rates of change in the structure of the total corpus callosum and sub-regions from term-equivalent age to 11 years of age, contrasted between treatment groups. Results are from separate mixed effects models for each corpus callosum region and parameter (area, thickness, diffusion value or tract volume), adjusted for sex of the child. The plots show the estimated means and 95% confidence intervals (CI) per group for each time point and the rate of change between time points from the mixed models. The y-axes show the parameters (area, thickness, diffusion value or tract volume) and the x-axes show age. Units are cm for thickness, cm^2 for area, cm^3 for volumes, $\times 10^{-3}$ mm^2/sec for diffusivities and years for age. The results are based on the total number of children who had usable MRI data at either term-equivalent age or 11 years of age; for the area measures, n = 144 (28 with term-equivalent data plus 116 with 11-year data); for the total tract diffusion and volume measures, n = 129 (20 plus 109); for the genu diffusion and volume measures, n = 137 (28 plus 109); for the rostral body diffusion and volume measures, n = 135 (26 plus 109); for the anterior mid body diffusion and volume measures, n = 133 (24 plus 109); for the posterior mid body diffusion and volume measures, n = 134 (25 plus 109); for the isthmus diffusion and volume measures, n = 134 (25 plus 109); for the splenium diffusion and volume measures, n = 137 (28 plus 109). FA, fractional anisotropy; AD, axial diffusivity; RD, radial diffusivity; MD, mean diffusivity.

decreases in diffusivities over time in the caffeine group may suggest that any early benefits of caffeine treatment on WM microstructure weaken over time.

However, caffeine appeared to influence corpus callosum size, and microstructure to a lesser extent, at age 11 years. The smaller fiber cross-section, thickness and

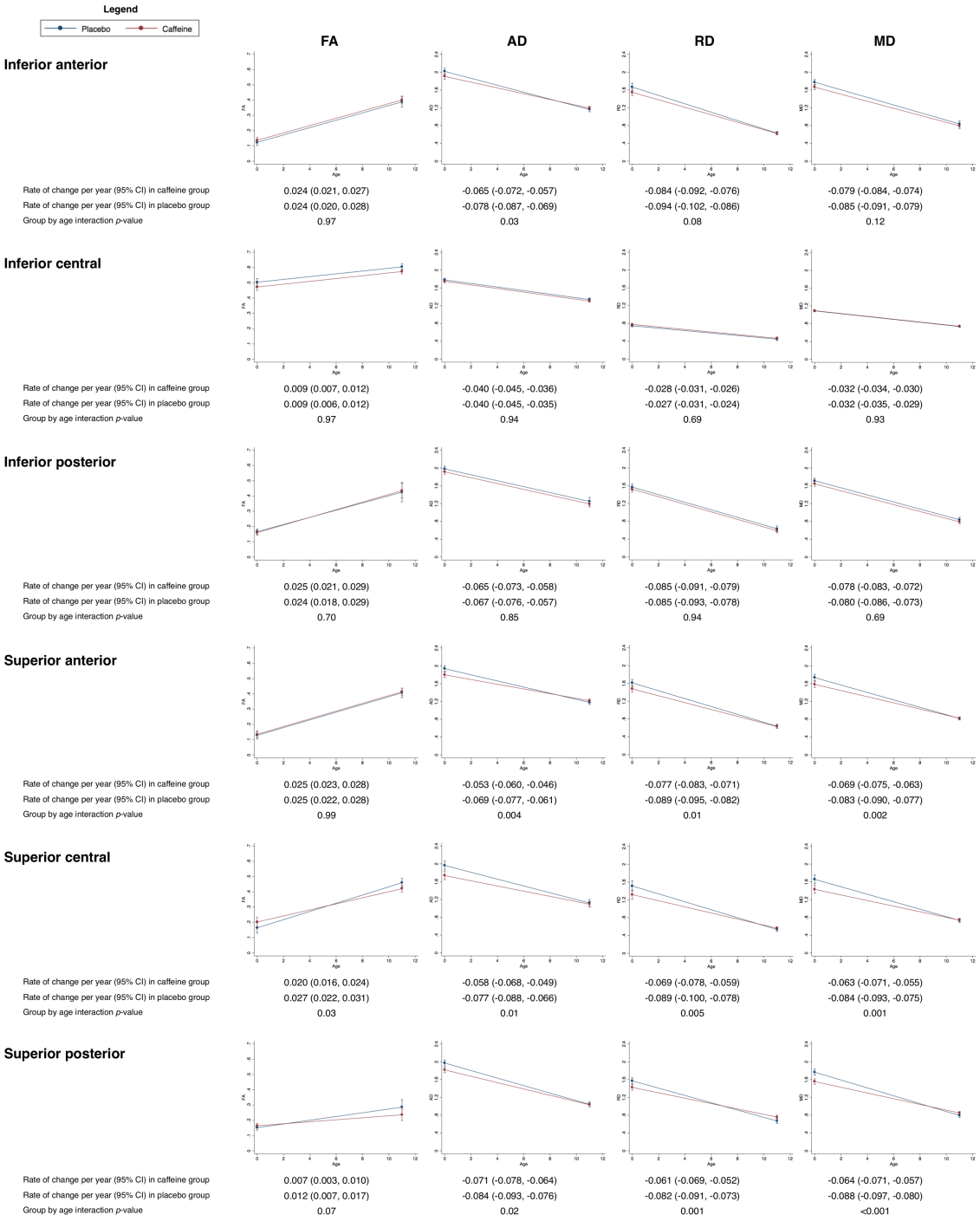


Figure 7. Rates of change in diffusion values in regions of interest from term-equivalent age to 11 years of age, contrasted between treatment groups. Results are from separate mixed effects models for each region and parameter (FA, AD, RD, and MD), adjusted for sex of the child. The plots show the estimated means and 95% confidence intervals (CI) per group for each time point and the rate of change between time points from the mixed models. The y-axes show the parameters (FA, AD, RD, MD) and the x-axes show age. Results are combined across brain hemispheres as there was negligible evidence that results differed by hemisphere. Units are $\times 10^{-3}$ mm^2/sec for diffusivities and years for age. The regions of interest were drawn on diffusion images from 28 infants at term-equivalent age.[7] Of those 28 infants, 19 had MRI at 11 years of age, and matching regions of interest were drawn on the 11-year diffusion images of these 19 children. These longitudinal analysis results are based on the 28 infants with usable term-equivalent data and the 19 children with usable 11-year data. FA, fractional anisotropy; AD, axial diffusivity, RD, radial diffusivity; MD, mean diffusivity.

area in the caffeine group compared with the placebo group found with both the fixel-based analysis and specific corpus callosum segmentation suggests the caffeine group's corpus callosum fibers are taking up less space. The magnitude of the area difference (~half a square centimeter) did not appear to be marginal, rather it is similar to the magnitude of the difference in corpus callosum area between very preterm and term-born children reported previously.[34] This finding could reflect axon loss in the caffeine group, with the concurrent lack of differences in the fiber density measure from the fixel-based analysis suggesting the remaining axons are still densely packed.[22] The finding could also reflect differences in myelin between caffeine and placebo groups, but we cannot be certain of this, because none of the techniques we used specifically measure myelin. Future analyses using MRI acquisitions and analysis techniques more sensitive to myelin would be beneficial.[35,36] Another possibility is that the finding reflects more coherent alignment of axons in the caffeine group. Because the precise cellular interpretation is unclear, it is not possible for us to conclude whether this difference would affect the ability of the corpus callosum to transfer information and in turn affect the cognitive or motor functioning of children treated with caffeine. Our finding may line up with some studies using animal models that reported adverse effects of caffeine on the developing brain.[6] Regardless, the clinical impact of this finding is likely to be negligible, given the overall CAP trial clearly indicated that caffeine is associated with benefits, and no harm, to short- and long-term health and neurodevelopmental outcomes for infants born low birthweight.[1–5]

The mechanisms for caffeine's beneficial neurological effects have never been clear.[2] Given that caffeine had no clear benefits on brain structure at age 11 years, our results do not support that brain structure at age 11 years provides a link between caffeine and improved neurodevelopmental outcomes, although brain-behavior relationships require direct investigation in future. The early improvements to WM microstructure in our previous study[7] may be more important than later brain development for determining neurodevelopmental outcomes. It is also possible that, rather than acting directly on the brain, caffeine may act indirectly by reducing comorbidities such as bronchopulmonary dysplasia and retinopathy of prematurity,[1,2] which themselves are strongly associated with neurodevelopmental disabilities.[37,38]

The major strength of our study is that it is part of a large randomized controlled trial, which is unlikely to be replicated due to the clear clinical benefits of caffeine. However, our study is limited in that only a small sample of children from the CAP trial had MRI, reducing our power to make inferences about the effects of caffeine on the brain. MRI is also inherently limited in that it does not directly measure brain structure and has too low resolution to identify alterations induced by caffeine at a cellular level. Additional follow-up of the CAP children would be beneficial, because until brain development has plateaued in late adolescence or adulthood, any differences between treatment groups may not reflect benefits or risks of caffeine, but may simply reflect normal variations during a period of rapid brain development. There are many factors known to influence brain development, which could confound our findings, including perinatal medical factors, environmental and sociodemographic factors, and pubertal factors.[30,31,39] Our results were independent of three important factors that influence brain development in preterm-born children: bronchopulmonary dysplasia, major neonatal brain injuries and maternal education level, which may relate to socioeconomic status.

In conclusion, our study suggests there are no clear benefits of caffeine on brain structure and development at age 11 years that are detectable by MRI, although caffeine may affect long-term corpus callosum development.

Acknowledgments

This research was conducted within the Developmental Imaging research group, Murdoch Children's Research Institute and the Children's MRI Centre, Royal Children's Hospital, Melbourne, Australia. We thank the Royal Children's Hospital Medical Imaging staff for their assistance and expertise in the collection of the MRI data included in this study. We thank the CAP Steering Committee, chaired by Barbara Schmidt, for permitting the release of the CAP study randomization sequence for the children who are included in this report. This work was supported by the National Health and Medical Research Council of Australia (Project Grants No. 237117 and No. 108706; Program Grant No. 606789; Centre of Research Excellence No. 1060733; Fellowship No. 1081288 to PJA, No. 1127984 to KJL, No. 1141354 to JLYC and No. 1085754 to DKT). This work was also generously supported by The Royal Children's Hospital Foundation (Grant no. RCH1000) devoted to raising funds for research at The Royal Children's Hospital, as well as the Murdoch Children's Research Institute, the Royal Children's Hospital, Department of Paediatrics at The University of Melbourne and the Victorian Government's Operational Infrastructure Support Program.

Role of the funding sources: The funding sources had no role in the design, implementation, analysis, interpretation, or writing of the results of this study, or in the decision to submit.

Author Contributions

Conception and design of the study: LWD, PJA, JLYC, DKT. Acquisition and analysis of data: all authors. Drafting a significant portion of the manuscript or figures (i.e., a substantial contribution beyond copy editing and approval of the final draft, which is expected of all authors): CEK.

Conflict of Interest

Nothing to report.

References

1. Schmidt B, Roberts RS, Davis P, et al. Caffeine therapy for apnea of prematurity. N Engl J Med 2006;354:2112–2121.
2. Schmidt B, Roberts RS, Davis P, et al. Long-term effects of caffeine therapy for apnea of prematurity. N Engl J Med 2007;357:1893–1902.
3. Doyle LW, Schmidt B, Anderson PJ, et al. Reduction in developmental coordination disorder with neonatal caffeine therapy. J Pediatr 2014;165(356–359):e352.
4. Schmidt B, Anderson PJ, Doyle LW, et al. Survival without disability to age 5 years after neonatal caffeine therapy for apnea of prematurity. JAMA 2012;307:275–282.
5. Schmidt B, Roberts RS, Anderson PJ, et al. Academic performance, motor function, and behavior 11 years after neonatal caffeine citrate therapy for apnea of prematurity: an 11-year follow-up of the CAP randomized clinical trial. JAMA Pediatr 2017;171:564–572.
6. Atik A, Harding R, De Matteo R, et al. Caffeine for apnea of prematurity: effects on the developing brain. Neurotoxicology 2017;58:94–102.
7. Doyle LW, Cheong J, Hunt RW, et al. Caffeine and brain development in very preterm infants. Ann Neurol 2010;68:734–742.
8. Ashburner J, Friston KJ. Unified segmentation. NeuroImage 2005;26:839–851.
9. Fischl B. FreeSurfer. NeuroImage 2012;62:774–781.
10. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis I. Segmentation and surface reconstruction. NeuroImage 1999;9:179–194.
11. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. NeuroImage 1999;9:195–207.
12. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341–355.
13. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 2006;31:968–980.
14. Patenaude B, Smith SM, Kennedy DN, et al. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 2011;56:907–922.
15. Diedrichsen J. A spatially unbiased atlas template of the human cerebellum. NeuroImage 2006;33:127–138.
16. Diedrichsen J, Balsters JH, Flavell J, et al. A probabilistic MR atlas of the human cerebellum. NeuroImage 2009;46:39–46.
17. Adamson C, Beare R, Walterfang M, et al. Software pipeline for midsagittal corpus callosum thickness profile processing : automated segmentation, manual editor, thickness profile generator, group-wise statistical comparison and results display. Neuroinformatics 2014;12:595–614.
18. Adamson CL, Wood AG, Chen J, et al. Thickness profile generation for the corpus callosum using Laplace's equation. Hum Brain Mapp 2011;32:2131–2140.
19. Thompson DK, Inder TE, Faggian N, et al. Characterization of the corpus callosum in very preterm and full-term infants utilizing MRI. NeuroImage 2011;55:479–490.
20. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 2006;31:1487–1505.
21. Bach M, Laun FB, Leemans A, et al. Methodological considerations on tract-based spatial statistics (TBSS). NeuroImage 2014. https://doi.org/10.1016/j.neuroimage.2014.06.021.
22. Raffelt DA, Tournier JD, Smith RE, et al. Investigating white matter fibre density and morphology using fixel-based analysis. NeuroImage 2017;144:58–73.
23. Beare RJ, Chen J, Kelly CE, et al. Neonatal brain tissue classification with morphological adaptation and unified segmentation. Front Neuroinform 2016;. https://doi.org/10.3389/fninf.2016.0001212.
24. Loh WY, Connelly A, Cheong JL, et al. A New MRI-Based Pediatric Subcortical Segmentation Technique (PSST). Neuroinformatics 2016;14:69–81.
25. Thompson DK, Wood SJ, Doyle LW, et al. Neonate hippocampal volumes: prematurity, perinatal predictors, and 2-year outcome. Ann Neurol 2008;63:642–651.
26. Greve DN, Van der Haegen L, Cai Q, et al. A surface-based analysis of language lateralization and cortical asymmetry. J Cogn Neurosci 2013;25:1477–1492.
27. Winkler AM, Ridgway GR, Webster MA, et al. Permutation inference for the general linear model. NeuroImage 2014;. https://doi.org/10.1016/j.neuroimage.2014.01.060.
28. Raffelt DA, Smith RE, Ridgway GR, et al. Connectivity-based fixel enhancement: whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. NeuroImage 2015;117:40–55.
29. Alexandrou G, Martensson G, Skiold B, et al. White matter microstructure is influenced by extremely preterm

birth and neonatal respiratory factors. Acta Paediatr 2014;103:48–56.
30. Inder TE, Warfield SK, Wang H, et al. Abnormal cerebral structure is present at term in premature infants. Pediatrics 2005;115:286–294.
31. Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat Rev Neurosci 2010;11:651–659.
32. Lebel C, Beaulieu C. Longitudinal development of human brain wiring continues from childhood into adulthood. J Neurosci 2011;31:10937–10947.
33. Thompson DK, Lee KJ, Egan GF, et al. Regional white matter microstructure in very preterm infants: predictors and 7 year outcomes. Cortex 2014;. https://doi.org/10.1016/j.cortex.2013.11.010.
34. Thompson DK, Lee KJ, van Bijnen L, et al. Accelerated corpus callosum development in prematurity predicts improved outcome. Hum Brain Mapp 2015;36:3733–3748.
35. Deoni SC, Rutt BK, Arun T, et al. Gleaning multicomponent T1 and T2 information from steady-state imaging data. Magn Reson Med 2008;60:1372–1387.
36. Glasser MF, Van Essen DC. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. J Neurosci 2011;31:11597–11616.
37. Anderson PJ, Doyle LW. Neurodevelopmental outcome of bronchopulmonary dysplasia. Semin Perinatol 2006;30:227–232.
38. Schmidt B, Asztalos EV, Roberts RS, et al. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months: results from the trial of indomethacin prophylaxis in preterms. JAMA 2003;289:1124–1129.
39. Lebel C, Deoni S. The development of brain white matter microstructure. NeuroImage 2018. https://doi.org/10.1016/j.neuroimage.2017.12.097.
40. Papile LA, Burstein J, Burstein R, et al. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr 1978;92:529–534.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Perinatal characteristics contrasted between included participants and excluded participants from the Royal Women's Hospital, Melbourne.
Table S2. Results of the fixel-based analysis, before and after adjusting for possible confounding factors.

