Neonatal white matter tract microstructure and 2-year language outcomes after preterm birth

Authors: Sarah E. Dubner, MD,[1] Jessica Rose, PhD,[2] Heidi M. Feldman, MD PhD,[1] Katherine E. Travis, PhD[1]

Affiliations:

[1]Division of Developmental-Behavioral Pediatrics, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States.

[2] Division of Pediatric Orthopaedics, Stanford University School of Medicine, Stanford, California

Corresponding Author:

Sarah E. Dubner, MD

sdubner@stanford.edu

Abstract:

Aim: To determine associations between 2-year language outcomes and infant diffusion MRI (dMRI) white matter tract metrics obtained prior to neonatal intensive care discharge in a cohort of preterm children. Method: 30 children, gestational age 28.9 (2.4) weeks, underwent dMRI at mean post menstrual age 36.4 (1.4) weeks and language assessment with the Bayley Scales of Infant Development–III at 22.2 (1.7) months chronological age. Mean fractional anisotropy (FA) and mean diffusivity (MD) were calculated for 5 white matter tracts. Hierarchical linear regression assessed associations between tract FA, moderating variables, and language. Results: Composite language correlated positively with FA of the left inferior longitudinal fasciculus (r = 0.44, (95% CI 0.03 – 0.71)) and negatively with FA of the posterior corpus callosum (r = -0.42, (95% CI -0.69, -0.06)). Number of neonatal medical complications moderated the relationship between language and posterior corpus callosum FA but did not moderate the association in the other tracts. Conclusion: Microstructural properties of white matter tracts in the newborn period were moderately associated with language outcomes in the toddler years. The different pattern of associations by fiber group may relate to the stage of brain maturation and/or the nature and timing of medical complications related to preterm birth. Future studies should replicate these findings with larger sample sizes to assure reliability of the findings.

Keywords: white matter; diffusion magnetic resonance imaging; language; infant, premature; tractography

1. INTRODUCTION[1]

Children born preterm (PT) are at risk for expressive and receptive language deficits (Barre et al., 2011; Foster-Cohen et al., 2007; van Noort-van der Spek et al., 2012). Language deficits in PT children increase risk for later reading problems, (Wocadlo and Rieger, 2007) which, in turn, are associated with adolescent health behaviors and long-term academic and occupational success (Connor et al., 2014; Sanders et al., 2009). Efforts to understand the neurobiological basis for early language development in PT children may suggest avenues of intervention to improve language skills.

PT infants are at risk for white matter alterations (Volpe, 2009). Between 24 and 32 weeks gestational age (GA), pre-oligodendrocytes, the immature progenitors of myelinating glial cells (ie.,oligodendrocytes), are susceptible to oxidative stress from hypoxia, ischemia, and inflammation, often failing to mature or myelinate properly (Volpe, 2009). Periventricular regions are particularly sensitive to oxidative stress-related injuries (Volpe, 2009). Diffusion MRI (dMRI) has become a preferred method for assessing and characterizing the properties of white matter pathways in the brain (Basser and Pierpaoli, 1996; Feldman et al., 2010). Common dMRI metrics used to

---

[1] Abbreviations: AD, axial diffusivity; ANOVA, one-way analysis of variance; AFQ, Automated Fiber Quantification; Arc-L, left arcuate fasciculus; Arc-R, right arcuate fasciculus; BSID-III, Bayley Scales of Infant Development, 3rd Edition; CI, Confidence Interval; CC-Occ, occipital segment of the corpus callosum; dMRI, diffusion magnetic resonance imaging; DTI, Diffusion Tensor Imaging; FA, fractional anisotropy; FT, full term; GA, gestational age; ILF-L, left inferior longitudinal fasciculus; ILF-R, right inferior longitudinal fasciculus; LPCH, Lucile Packard Children's Hospital; MD, mean diffusivity; MedRisk, number of medical complications; MRI, magnetic resonance imaging; NICU, Neonatal Intensive Care Unit; PMA, post menstrual age; PT, preterm; RD, radial diffusivity; ROI, region of Interest; TEA, term equivalent age; UF-L, left uncinate fasciculus; UF-R, Right uncinate fasciculus; VIF, Variance Inflation Factor

index white matter microstructure are fractional anisotropy (FA) and mean diffusivity (MD). Within white matter voxels, increased FA and decreased MD are generally associated with favorable neurobiological factors, such as increased myelination and axonal density. However, dMRI metrics are also sensitive to additional structural and tissue properties that reduce FA and/or increase MD, including the proportion of crossing fibers within a voxel, axonal size, and membrane permeability (Jones and Cercignani, 2010). Evidence for lower FA and/or higher MD in dMRI studies of PT neonates performed at term equivalent age (TEA) compared with children born full term (FT) (Ball et al., 2013; Thompson et al., 2011) is consistent with evidence for reduced myelination and axonal growth in animal models of PT birth (Volpe, 2009).

Studies have examined relations between language and white matter pathway metrics in school-aged PT and FT children (Dodson et al., 2018; Travis et al., 2016), but few studies have examined language in relation to white matter metrics in the neonatal period, before children leave the neonatal intensive care unit (NICU), which usually occurs near 35 – 45 weeks post-menstrual age, or PMA). Emerging evidence suggests that variability in white matter microstructure obtained with dMRI prior to NICU discharge relates to later language development in PT children. FA in the left and right arcuate fasciculi measured at 39-46 weeks in children born preterm was found to be independently associated with language outcomes at 22 months (Salvan et al., 2017). In a longitudinal cohort of PT children imaged at TEA, 2- and 4-years of age, (Young et al., 2017) slower rate of change of MD and radial diffusivity (RD) in the left internal and external capsules, and slower rate of change of axial diffusivity (AD) in the left posterior

thalamic radiation was associated with lower 4-year old language scores. Additional studies are required to confirm and expand on these associations.

PT birth-related illnesses, such as bronchopulmonary dysplasia, have also been associated with white matter metrics (Dubner et al., 2019; Shim et al., 2012) and with language outcomes (Singer et al., 2001). The degree to which PT children experience neonatal medical complications varies substantially. Studies have yet to determine whether medical complications of PT birth, especially those likely to induce hypoxic, ischemic or neuroinflammatory responses, moderate associations between pre-discharge measures of white matter microstructure and language outcomes. Examining these relations in detail is pertinent to determining whether white matter microstructure may be an important early biomarker for language outcomes.

The present study used advanced, anatomically specific tractography methods to determine if language outcomes at 2 years related to variability in white matter microstructure across multiple language-related white matter tracts on pre-discharge infant dMRI scans in PT children and whether the number of medical complications moderated any effects. We obtained measures for FA and MD from specific white matter pathways commonly implicated in processing linguistic information in infants and young children, including a dorsal pathway, the left arcuate fasciculus (Arc-L) and two bilateral ventral pathways, the left and right inferior longitudinal fasciculi (ILF-L and ILF-R) and the left and right uncinate fasciculi (UF-L, UF-R) (Hickok and Poeppel, 2007). Additionally, we examined fiber pathways traversing the occipital segment of the corpus callosum (CC-Occ) for several reasons. Intrahemispheric connections are postulated to play an important role in language processing in infancy (Perani et al., 2011). Studies

have found relations between language-related skills and the posterior corpus callosum in both PT (Dubner et al., 2019) and FT children (Dougherty et al., 2007; Frye et al., 2008; Hasan et al., 2012; Huber et al., 2019; Odegard et al., 2009). In addition, this region is susceptible to injuries from medical complications of PT birth (Back, 2015; Volpe, 2009). We focused our analyses on FA and MD given our previous work, which found these metrics to be related to neurocognitive outcomes in sample of older children born PT with and without history of neonatal inflammation (Dubner et al., 2019). We hypothesized that language would be associated with white matter metrics in all of the selected tracts, independent of PMA. In addition, we hypothesized that the number of neonatal medical complications would moderate the association in posterior white matter tracts because of their putative susceptibility to oxidative-stress related injuries.

2. MATERIALS AND METHODS

2.1 Participants

Children ≤ 1500 grams or ≤ 32 weeks gestation at birth from the Lucile Packard Children's Hospital (LPCH) Stanford Neonatal Intensive Care Unit (NICU) born between January 2010 and December 2011 were eligible for this prospective study (Rose et al., 2015). Exclusion criteria included evidence of genetic disorders or congenital brain anomalies. 66 preterm infants underwent dMRI imaging near hospital discharge. Per LPCH NICU protocol, MRIs were obtained when infants were stable in an open crib, requiring no more than supplemental oxygen for respiratory support, and > 34 weeks PMA. Of these, n = 36 participants were excluded due to T1-weighted scan issues (n = 9), incomplete diffusion scans (n = 1), excessive head movement (n = 21), presence of

parenchymal brain disease, including periventricular leukomalacia, on clinical MRI (n = 2), or missing language assessment (n = 3). The final sample was 30 participants. There were no significant differences between the 30 included and 36 excluded participants from the initial sample in GA (28.9 (2.4) vs 28.9 (2.4) weeks, t(64) = 0.05, *p* = 0.96), birthweight (1085 (281) vs 1095 (258) grams, t(64) = 0.14, *p* = 0.89), or PMA (36.4 (1.4) vs 36.6 (1.2) weeks, t(64) = 0.72, *p* = 0.49). In addition there were no differences between included and excluded participants in proportion of males (36.7% vs 38.9%, *p* = 0.85), or proportion with sepsis (3.3% vs 13.9%, *p* = 0.21), bronchopulmonary dysplasia (16.7% vs 33.3%, *p* = 0.16), necrotizing enterocolitis (3.3% vs 16.8%, *p* = 0.12), patent ductus arteriosus (48.3% vs 41.7%, *p* = 0.62), retinopathy of prematurity (36.7% vs 41.7%, *p* = 0.68), or intraventricular hemorrhage (20.0% vs 22.2%, *p* = 0.83).

Clinical variables were obtained via chart review. Additional description of the participants and outcomes using different analytic strategies was previously reported (Rose et al., 2015). The experimental protocol was approved by the Stanford University Institutional Review Board #IRB-13899. A parent or legal guardian provided informed written consent and participants were compensated for participation in neurodevelopmental assessments.

2.2 Procedures

2.2.1 Diffusion MRI Acquisition, Measures, and Analyses

2.2.1.1 MRI ACQUISITION

MRI data were acquired on a 3T Discovery MR750 scanner equipped with an 8-channel HD head coil (General Electric Healthcare, Little Chalfont, UK). Infants were swaddled and fed and typically remained asleep during the scan. Sedation was not utilized as part of the research protocol and was typically not utilized for routine MRI. Clinical neuroradiological assessment of the conventional MRI scans were obtained by chart review. High-resolution T1-weighted anatomical images were collected for each subject. A gradient echo 3-plane localizer was used, and an asset calibration was proscribed to utilize parallel imaging. From the 3-plane localizer, T1-weighted images were collected with the following parameters: TE = 91.0, TR = 2200, FOV = 20 cm, matrix size = 320 × 224, slice thickness 3.0 × 0.5 mm spacing, NEX = 1. This T1-weighted image was used as a common reference for alignment of the diffusion tensor image (DTI) maps. dMRI data were acquired with a single-shot, spin- echo, echo-planar imaging sequence with a slice thickness of 3.0 mm, a matrix size of 128 × 128, a 90° flip angle, FOV = 20 cm, TE = 88.8 ms, TR = 8,000, (b = 1,000 s/mm^2). Diffusion was measured along 25 directions, with three b=0 images.

2.2.1.2 DATA PREPROCESSING

Data preprocessing steps, motion correction procedures, diffusion tensor estimation, and individual native space tractography are described below and have been described in our previous work (Dodson et al., 2017; Travis et al., 2017). The T1-weighted images were first aligned to the canonical ac-pc orientation. Diffusion weighted images were pre-processed with open-source software, mrDiffusion (*http://white.stanford.edu/newlm/index.php/MrDiffusion*) implemented in MATLAB R2012a (Mathworks, Natick, MA). Subjects' motion during the diffusion-weighted scan

was corrected using a rigid body alignment algorithm (Rohde et al., 2004). Each diffusion weighted image was registered to the mean of the three non-diffusion (b=0) images and the mean b=0 image was registered automatically to the participant's T1-weighted image, using a rigid body transformation (implemented in SPM8, http://www.fil.ion.ucl.ac.uk/spm/; no warping was applied). The combined transform that resulted from motion correction and alignment to the T1 anatomy was applied to the raw data once, and the transformed images were resampled to 2 x 2 x 2 mm isotropic voxels. This step was performed because non-isotropic voxels may bias the tensor fit and distort both tracking and measurements of diffusion properties (Oouchi et al., 2007). Diffusion gradient directions were then adjusted to fit the resampled diffusion data (Leemans and Jones, 2009).

For each voxel in the aligned and resampled volume, tensors were fit to the diffusion measurements using a standard least-squares algorithm, Robust Estimation of Tensors by Outlier Rejection (RESTORE), which is designed to remove outliers at the tensor estimation step (Chang et al., 2005). A continuous tensor field was estimated using trilinear interpolation of the tensor elements. The eigenvalue decomposition of the diffusion tensor was calculated and the resulting three eigenvalues ($\lambda 1$, $\lambda 2$, $\lambda 3$) were used to compute FA, MD (i.e., the mean of $\lambda 1$, $\lambda 2$, and $\lambda 3$), AD ($\lambda 1$), and RD (($\lambda 2$ + $\lambda 3$)/2) (Basser and Pierpaoli, 1996).

2.2.1.3 HEAD MOTION

All dMRI scans were rigorously assessed for head motion using the following procedures. We quantified the amount of relative head motion (in mm) in each participant by calculating the amount of motion correction (in mm) in the x-y-z plane of

each volume relative to the prior volume in the sequence. For each diffusion-weighted scan, we quantified the degree of relative head motion in each participant by calculating the magnitude of motion correction in voxels (voxel size = 3 mm) in the x-y-z planes of each volume relative to the prior volume. We then calculated the mean relative motion for each participant. Participants who had mean relative motion greater than 0.7 mm were excluded (Huber et al., 2019; Yendiki et al., 2014). Using these stringent criteria, 21 of the 66 participants were excluded.

2.2.1.4 TRACTOGRAPHY

Current analyses focused on *a priori selected* white matter tracts that have been previously implicated in processing linguistic information, depicted in **Figure 1.** Specifically, we identified tracts of a dorsal pathway, the Arc-L and two bilateral ventral pathways, the ILF-L and ILF-R, and the UF-L and UF-R. Additionally, we examined the CC-Occ to capture posterior white matter regions associated with language related skills (Dougherty et al., 2007; Dubner et al., 2019; Frye et al., 2008; Hasan et al., 2012; Huber et al., 2019; Odegard et al., 2009) and susceptible to injuries from PT-birth related complications (Back, 2015; Volpe, 2009). Due to previously described limitations of deterministic tractography for segmenting the right arcuate fasciculus, we made the *a priori* decision to exclude the right arcuate fasciculus from our analyses (Catani et al., 2007; Lebel and Beaulieu, 2009; Mishra et al., 2010; Travis et al., 2015a; Yeatman et al., 2011).

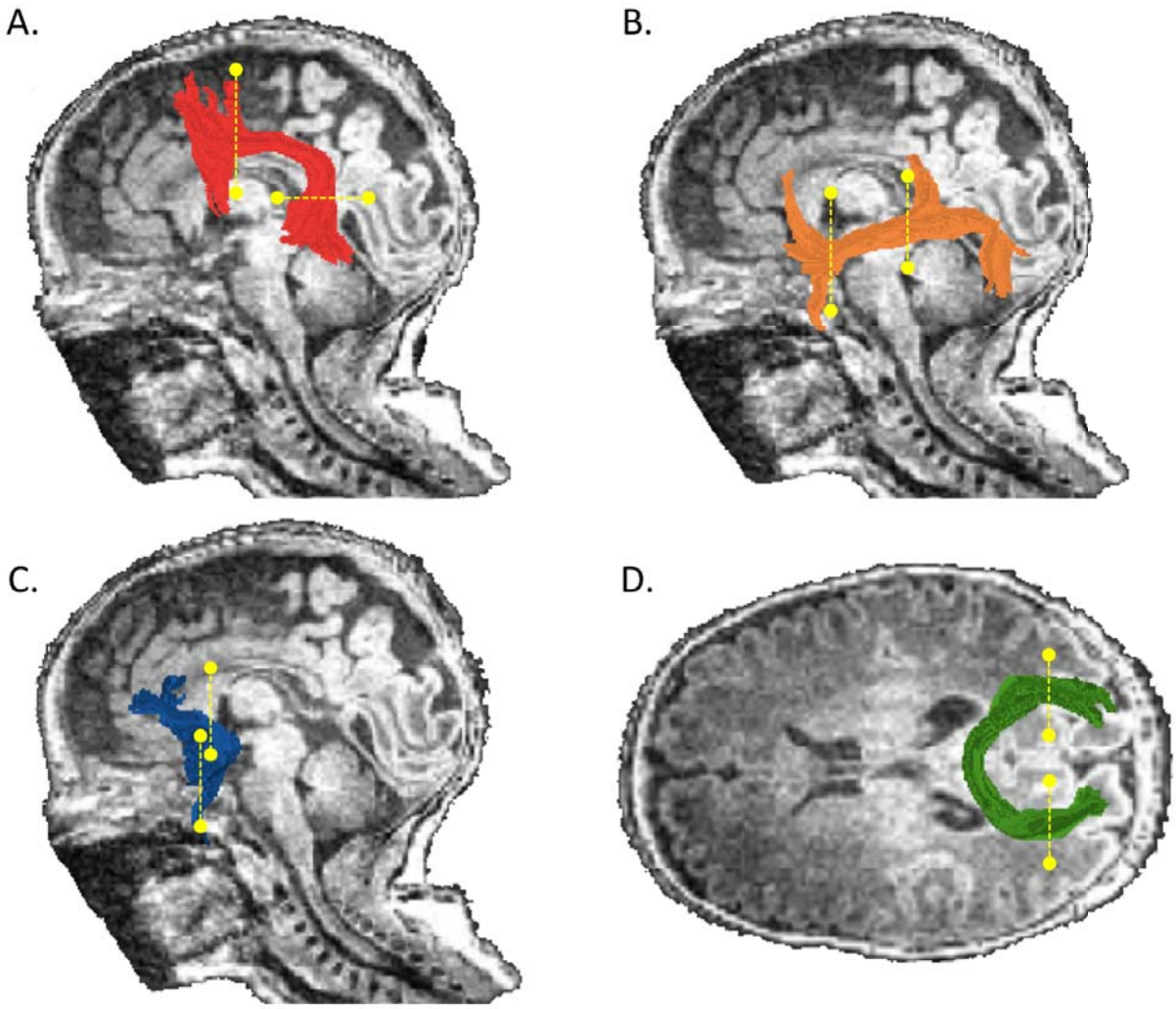

Figure 1. Tractography of selected white matter tracts. Left hemisphere tract renderings are displayed on a mid- sagittal T1 image from a representative participant. Right hemisphere tract renderings not shown. A.) The arcuate fasciculus is shown in red, B.) the inferior longitudinal fasciculus is shown in orange, C.) the uncinate fasciculus is shown in blue. D.) Occipital segment of the corpus callosum rendering is displayed in green on an axial T1 image from the same participant. Dashed lines represent the location of the regions of interest (ROIs) used to segment each pathway from the whole-brain tractogram.

In order to maximize sensitivity while taking into account considerable individual variability, particularly at this early stage in brain development, our approach used individual tract identification in the native space of each child, followed by quantification of diffusivity properties along the length of the tract. We achieved this procedure by performing automatic tract segmentation and quantification using the Automated Fiber Quantification (AFQ; https://github.com/yeatmanlab/AFQ) software package and MATLAB. AFQ uses a three-step procedure to identify each tract in each individual: whole brain tracking, automatic segmentation using anatomically defined template ROIs in MNI space warped to native space, and automatic refinement and cleaning (Yeatman et al., 2012b). The AFQ pipeline has been successfully implemented for neonatal images in infants as young as 1 day old (Bruckert et al., 2019; Langer et al., 2017). AFQ procedures have been described extensively in several previous studies (Dodson et al., 2017; Travis et al., 2015a, 2015b, 2016, 2017; Yeatman et al., 2012b). In the current study, our approach for segmenting dorsal and ventral tracts was similar to Travis *et al*., (Travis et al., 2015a). Procedures for segmenting the posterior corpus callosum were similar to those described in (Dubner et al., 2019). Briefly, this procedure involves segmenting the corpus callosum into 8 discrete and non-overlapping regions based on the cortical zone where fiber projections terminated (ie., occipital terminations for the posterior corpus callosum), a method that allows for anatomically specific and functionally relevant divisions of the corpus callosum (Dougherty et al., 2007; Huang et al., 2005).

To account for the immaturity of the infant brain, we adjusted the fiber tracking algorithm to perform tracking beginning in voxels from a white matter mask of FA > 0.1

and to stop tracking when voxels dropped below a threshold of FA = 0.08. Individual tractograms representing each fiber group in each subject were visually inspected to confirm that each tract conformed to anatomical norms. We performed this step to ensure that fiber groups did not contain multiple additional streamlines or false positives because of the low threshold needed to perform tracking in the neonatal brain. Tracts that did not conform to anatomical norms were excluded from statistical analyses. After visual inspection, we retained 80% of ILF-L tracts, 73% of ILF-R tracts, 100% of UF-L tracts, 100% of UF-R tracts, 93% of CC-Occ tracts, and 47% of Arc-L tracts across participants. In 14 participants, the Arc-L contained large proportions streamlines that were not consistent with its known anatomy and that were not excluded via automated fiber segmentation procedures implemented here. Since the Arc-L conformed to anatomical norms in only (n=16) 53% of participants, we chose to exclude this tract from further analyses.

2.2.1.5 FIBER TRACT QUANTIFICATION: TRACT PROFILES

For each of the tracts, diffusion properties (FA, MD) were calculated at 30 equidistant nodes along a central portion of each fiber tract bounded by the same two ROIs used for tract segmentation. This procedure generated, for every tract and every individual, an FA tract profile that described the variations in FA along the central portion of the tract. At each node, diffusion properties were calculated by taking a weighted average across all streamlines belonging to this tract. Each streamline's contribution to the average was weighted by the probability that a fiber was a member of the fascicle, computed as the Mahalanobis distance from the tract core (Yeatman et al., 2012b). This procedure minimizes the contribution of fibers located further from the fiber

tract core that are more likely to reflect a mixture of gray and white matter or of different tracts, and thereby minimizes the effect of partial voluming on diffusion estimates. This procedure further ensures that we are able to characterize core tract regions that are anatomically consistent across individuals. Tract profiles were then averaged to produce a single mean value for each of the tracts.

2.2.2 Neurodevelopmental assessment

Participants attended the LPCH High-Risk Infant Follow-Up program, during which they were assessed on the Bayley Scales of Infant Development, 3rd Edition (BSID-III) (Bayley, 2006). The target age for assessment was 18-24 months corrected age. One participant who was assessed at 29.5 months chronological age was included because the use of standard scores adjusts for age. We analyzed the composite language standardized score and the expressive and receptive language subtest scaled scores.

2.2.3 Medical complications

To assess the contributions of number of medical complications to outcomes, we generated a composite variable, MedRisk, similar to that used in previous studies (Adams et al., 2018; Marchman et al., 2019, 2016). MedRisk was calculated by summing the presence of common medical complications of PT birth that increase the risk for poor long term cognitive or behavioral outcomes. These medical complications were intraventricular hemorrhage, bronchopulmonary dysplasia, necrotizing enterocolitis, blood culture proven sepsis, patent ductus arteriosus, and retinopathy of prematurity. No participant had seizures or hearing loss, and all but three had hyperbilirubinemia, therefore these common medical complications were excluded from

the MedRisk variable. Participants were grouped into low risk (0 risk factors), medium risk (1 or 2 risk factors), and high risk (≥ 3 risk factors).

2.3 Statistical analyses

Statistical analyses were conducted using SPSS (version 26.0.0.1, IBM Corp., 2019). Statistical significance was set at $p < 0.05$. The Shapiro–Wilk test was used to assess whether clinical and neurobiological data were normally distributed. With the exception of PMA, data were normally distributed. We chose to do parametric tests for all associations.

Statistical analyses were planned based on the initial sample of 66 participants. With $\alpha = 0.05$ and $1 - \beta = 0.8$, the study was powered to detect a correlation effect size of 0.34. Based on the final sample size of 30 participants with $\alpha = 0.05$ and $1 - \beta = 0.8$, the study was powered to detect a correlation effect size of 0.49.

2.3.1 Demographic and clinical variables

T-tests were computed to evaluate language outcomes for differences based on sex. Pearson correlations were computed to evaluate associations between language outcomes or white matter microstructure metrics, GA and PMA. One-way ANOVA assessed whether language outcomes differed by MedRisk.

2.3.2 Mean-tract associations between two-year language outcomes and infant FA and MD

Pearson correlation coefficients and 95% confidence intervals (CI) were computed to assess zero-order associations between composite, receptive, and expressive language scores and mean FA or MD. Though we chose the tracts of

interest based on previous data, to be conservative, a False Discovery Rate of 5% was calculated for zero-order associations to account for associations across 5 tracts (Benjamini and Hochberg, 1995).

We then conducted a series of hierarchical linear regression models to assess whether number of medical complications moderated associations between mean FA or MD and language outcomes. Given the modest sample size, we chose to restrict regression analyses to tracts found to demonstrate correlations with composite language scores in which the 95% CI did not include zero. PMA was entered in the models to ensure results were related to white matter microstructure metrics rather than participant age at scan. We assessed the main effects of MedRisk and the white matter metrics. To explore the pattern of associations across MedRisk, we next determined if MedRisk moderated the prediction of FA to language outcome (Hayes, 2017)**.** The variance inflation factor (VIF) was calculated to assess multicollinearity of each model. We considered VIF values less than 10 to indicate that there was no concern for multicollinearity (Field, 2013).

2.4 Secondary Analyses

2.4.1 Along-tract profile associations between two-year language outcomes and infant white matter metrics

We calculated Pearson correlations of the composite language outcome measure and FA or MD values at each location along the tract profile. This analysis was conducted in order to confirm that correlations between language and either FA or MD were not obscured due to the use of mean tract measures. The analysis of tract profiles

utilized a nonparametric permutation-based method to control for 30 comparisons along the tract (Nichols and Holmes, 2002). This procedure produced a critical node cluster size for each of the candidate tracts (significant cluster size for all tracts was ≥6 locations/nodes). Correlations along a tract were considered significant after correction for within tract comparisons if they occurred in a cluster greater than or equal to than the critical cluster size.

2.4.2 Mean-tract associations between language outcomes and infant AD and RD

We conducted secondary analyses to evaluate the contribution of AD and RD to the correlations between diffusion metrics (FA or MD) and two-year composite language outcomes. These analyses were performed only for those tracts in which 95% CI did not include zero. This step was achieved by calculating Pearson correlation coefficients and 95% CIs between composite language and the mean AD or mean RD.

3. RESULTS

3.1 Participant Characteristics

Participant characteristics, neonatal medical conditions, and language outcomes are shown in **Table 1.** PMA and GA were weakly negatively correlated (r = -0.186, *p* = 0.33). Six participants had intraventricular hemorrhage on head ultrasound, all of which was grade I. Clinical MRI scans demonstrated no abnormality (n = 16) or minimal abnormality (n = 14), such as subependymal hemorrhage or mineralization.

Table 1. Participant Characteristics (n=30)

Participant Demographics	**mean (sd)**	**min - max**
Gestational Age, weeks	28.9 (2.3)	24.9 – 32.9
Birthweight, grams	1085 (281)	520 - 1510
Post Menstrual Age at Scan, weeks	36.3 (1.4)	34.7 – 40.3
Chronological Age at Bayley, months[a]	22.2 (1.7)	20.6 – 29.4
Adjusted Age at Bayley, months[a]	19.5 (1.9)	18.1 – 27.3
Maternal Age, years	33.2 (6.5)	13 - 47
Male, n (%)	11 (36.7)	
Neonatal Medical Conditions		**n (%)**
Intraventricular Hemorrhage		6 (20.0)
Clinical MRI Findings		
No Abnormality		16 (50.0)
Minimal Abnormality		14 (43.8)
Bronchopulmonary Dysplasia		5 (16.7)
Sepsis		1 (3.3)
Necrotizing Enterocolitis		1 (3.3)
Patent Ductus Arteriosus		14 (46.7)
Retinopathy of Prematurity		11 (36.7)
Stage 1		7 (23.3)
Stage 2		3 (10.0)
Stage 3		1 (3.3)
Hyperbilirubinemia		27 (90)
Number of Medical Risk Factors[b]		
0		10 (33.3)
1		9 (30.0)
2		4 (13.3)
3		7 (23.3)
Bayley Scales of Infant Development III	**mean (sd)**	
Composite Language		83.9 (16.8)
Expressive Scaled Score		7.6 (3.1)
Receptive Scaled Score		7.5 (3.4)

[a] n = 27. Age at developmental testing was missing for 3 participants. [b] Medical Risk Factors include intraventricular hemorrhage, bronchopulmonary dysplasia, sepsis, necrotizing enterocolitis, patent ductus arteriosus, retinopathy of prematurity. All intraventricular hemorrhage was grade I.

3.2 Language outcomes at two years

Participants demonstrated below-average composite language scores with substantial variation. Language scores did not differ by sex (p = 0.46) or MedRisk ($F(3, 26) = 0.55$, p = 0.65), and were not significantly correlated with PMA at scan (r = 0.03, p

= 0.87) or GA (r = 0.09, *p* = 0.64).

3.3 Measures of white matter microstructure

Table 2 lists tracts, number of participants in which each tract was identified, and mean (SD) FA and MD. Correlations between tract FA or MD and PMA are shown in **Table 3**. FA of the ILF-L and ILF-R positively correlated with PMA, indicating that as expected FA increased with increasing PMA. MD of the ILF-R, UF-L, and UF-R negatively correlated with PMA, also the expected direction.

Table 2. Tract mean FA and MD

Tract	n	FA mean (SD)	MD mean (SD)
ILF-L	24	0.19 (0.02)	1.57 (0.09)
ILF-R	22	0.18 (0.02)	1.51 (0.07)
UF-L	30	0.19 (0.02)	1.40 (0.06)
UF-R	30	0.20 (0.02)	1.39 (0.05)
CC-Occ	28	0.33 (0.04)	1.56 (0.08)

FA = Fractional Anisotropy, MD = Mean Diffusivity (10^{-3} mm^2/sec). ILF = inferior longitudinal fasciculus, UF = uncinate fasciculus, CC-Occ = occipital segment of the corpus callosum. L = Left, R = Right

Table 3. Pearson correlations between PMA and tract mean FA and MD

Tract	n	r_{FA}	*p*	r_{MD}	*p*
ILF-L	24	0.70	$2x10^{-4}$	-0.17	0.44
ILF-R	22	0.82	$3x10^{-6}$	-0.67	$1x10^{-3}$
UF-L	30	0.32	0.08	-0.39	0.03
UF-R	30	0.25	0.19	-0.39	0.03
CC-Occ	28	0.07	0.72	-0.23	0.24

PMA = Post Menstrual Age at dMRI scan, FA = Fractional Anisotropy, MD = Mean Diffusivity, r = Pearson correlation, ILF = inferior longitudinal fasciculus, UF = uncinate fasciculus, CC = corpus callosum, Occ = occipital.

3.4 Mean-tract associations between language outcomes and FA and MD

Correlations between language scores and mean FA are shown in **Table 4A**. Composite language scores positively correlated with FA of the ILF-L, and negatively

correlated with FA of the CC-Occ. Receptive and expressive language scores positively correlated with FA of the ILF-L. The 95% CIs for the correlation coefficients for composite language with ILF-L FA and with CC-Occ did not include zero and were therefore selected for further analyses. Although these correlations were moderate in strength, none survived False Discovery Rate correction for multiple comparisons.

Correlations between language outcomes and mean MD are shown in **Table 4B**. Receptive language scores moderately positively correlated with MD of the UF-L and the 95% CI did not include zero. The correlation did not survive False Discovery Rate correction for multiple comparisons.

Table 4. Pearson correlations between tract mean FA (Panel A) and MD (Panel B) and language outcomes

A.

		FA								
		Composite Language			**Receptive Language**			**Expressive Language**		
Tract	**n**	**r**	***p***	**95% CI**	**r**	***p***	**95% CI**	**r**	***p***	**95% CI**
ILF-L	24	**0.44**	**0.03**	**(0.04, 0.71)**	**0.47**	**0.02**	**(0.08, 0.73)**	**0.45**	**0.03**	**(0.06, 0.72)**
ILF-R	22	0.12	0.60	(-0.32, 0.51)	-0.02	0.92	(-0.44, 0.40)	0.24	0.27	(-0.20, 0.60)
UF-L	30	-0.24	0.20	(-0.55, 0.13)	-0.17	0.36	(-0.50, 0.20)	-0.13	0.49	(-0.47, 0.24)
UF-R	30	-0.28	0.14	(-0.58, 0.09)	-0.27	0.14	(-0.57, 0.10)	-0.10	0.59	(-0.45, 0.27)
CC-Occ	28	**-0.42**	**0.03**	**(-0.69 -0.06)**	-0.34	0.07	(-0.63 0.03)	-0.28	0.15	(-0.59, 0.10)

B.

		MD								
		Composite Language			**Receptive Language**			**Expressive Language**		
Tract	**n**	**r**	***p***	**95% CI**	**r**	***p***	**95% CI**	**r**	***p***	**95% CI**
ILF-L	24	-0.19	0.36	(-0.55, 0.23)	-0.23	0.28	(-0.58, 0.19)	-0.19	0.36	(-0.55, 0.23)
ILF-R	22	-0.18	0.41	(-0.56, 0.26)	-0.05	0.81	(-0.47, 0.38)	-0.32	0.15	(-0.65, 0.12)
UF-L	30	0.34	0.06	(-0.02, 0.63)	**0.43**	**0.02**	**(0.08, 0.68)**	0.21	0.26	(-0.16, 0.53)
UF-R	30	0.07	0.71	(-0.30, 0.42)	0.06	0.74	(-0.30, 0.41)	0.03	0.88	(-0.33, 0.39)
CC-Occ	28	0.02	0.91	(-0.39, 0.35)	-0.05	0.78	(-0.42, 0.32)	-0.20	0.31	(-0.53, 0.19)

FA = Fractional Anisotropy, MD = Mean Diffusivity. ILF = inferior longitudinal fasciculus, UF = uncinate fasciculus, CC-Occ = occipital segment of the corpus callosum. L = Left, R = Right. CI = Confidence Interval. Correlations do not survive false discovery rate correction. Bolded values indicate correlations where the 95% CI does not include zero.

3.5 Hierarchical multivariable linear regression models

Table 5 shows the results of multivariable linear regression models predicting language outcome at age 2 from FA of ILF-L and CC-Occ, beyond variation accounted for by PMA and MedRisk. FA of the ILF-L contributed to composite language **(Model 1B)**, explaining an additional 16.8% of the variance above that accounted for by PMA and MedRisk alone **(Model 1A)**. Over 41% of the variance in language was accounted for by the model**.** The interaction term did not contribute additional variance (**Model 1C**), indicating that the relationship between language outcomes at age 2 and FA of the ILF-L in infancy was similar for children regardless of the number of MedRisk conditions (**Figure 2A**).

FA of the CC-Occ contributed to composite language (**Model 2B**), explaining an additional 19.4% of the variance above that accounted for by PMA and MedRisk alone **(Model 2A)**. Over 20% of the variance in language scores was accounted for by the model. The overall model was not significant. The interaction of CC-Occ and MedRisk increased the variance accounted for to 34.1% (**Model 2C**). Moderation analysis indicated that a negative association between language outcomes and FA of the CC-Occ in infancy was found in the participants with 1 or more MedRisk conditions, but not zero MedRisk conditions, as illustrated in **Figure 2B**.

Table 5. Prediction of language outcome at age 2 years from mean tract-FA of A. the left inferior longitudinal fasciculus mean tract-FA (ILF-L FA), and B**.** Corpus Callosum Occipital segment mean tract-FA (CC-Occ FA), beyond covariates of post menstrual age at scan (PMA Scan) and level of medical complications (MedRisk).

A.

	Model 1A (n = 24)	**Model 1B**	**Model 1C**
PMA Scan	0.65 (0.12, 1.17)	0.31 (-0.24, 0.87)	0.23 (-0.46, 0.92)
Med Risk	-0.53 (-1.05, -0.01)	-0.65 (-1.13, -0.16)	-1.42 (-5.32, 2.48)
ILF-L		0.59 (0.08, 1.09)	0.49 (-0.21, 1.19)
ILF-L FA x MedRisk			0.89 (-3.56, 5.34)
Δ R^2		0.17 (p = 0.03)	0.01 (p = 0.68)
Total R^2 (95% CI)	0.25	0.42 (0.17, 0.67) (p = 0.01)	0.42 (0.18, 0.66) (p = 0.03)
Adj R^2	0.18	0. 33	0.30

B.

	Model 2A (n = 28)	**Model 2B**	**Model 2C**
PMA Scan	0.05 (-0.38, 0.48)	0.10 (-0.30, 0.50)	0.02 (-0.35, 0.40)
Med Risk	-0.11 (-0.54, 0.32)	-0.18 (-0.57, 0.22)	3.49 (-0.02, 7.00)
CC-Occ FA		-0.45 (-0.83, -0.07)	0.16 (-0.52, 0.83)
CC-Occ FA x MedRisk			-3.65 (-7.12, -0.18)
Δ R^2		0.19 (p=0.02)	0.14 (p=0.04)
Total R^2 (95% CI)	0.01	0.21 (-0.02, 0.44) (p=0.13)	0.34 (0.11, 0.58) (p=0.04)
Adj R^2	-0.07	0.11	0.23

Data are standardized coefficients (95% Confidence Interval). R^2 change (**Δ** R^2) values in models B and C, are in reference to model A. **Δ** R^2 values in models B and C reflect the increase in variance accounted for by the interaction term in relation to the preceding model with the main effect for that tract. MedRisk is a composite variable from 0 – 6 calculated as the sum of the following medical risk factors: necrotizing enterocolitis, intraventricular hemorrhage, bronchopulmonary dysplasia, patent ductus arteriosus, retinopathy of prematurity, and culture proven sepsis.

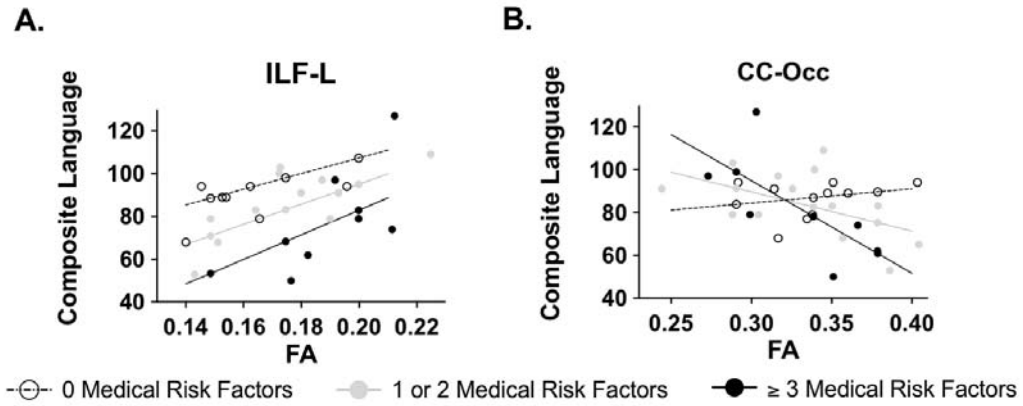

Figure 2. Associations of FA from infant scans obtained prior to NICU discharge and Bayley Scales of Infant Development, 3rd Edition Composite Language Scores, obtained at 18-29 months. Correlations are visualized as scatter plots between mean FA values and composite language values. Predicted model, which is adjusted for post menstrual age at scan, is visualized as regression lines for each MedRisk level. Children with zero medical risk factors are represented with open circles and dashed lines, children with one or two medical risk factors are represented with closed gray circles and solid gray lines, and children with three or more medical risk factors are represented with closed black circles and solid black lines. A) In the left inferior longitudinal fasciculus (ILF-L) the degree of association of ILF-L FA and language was the same regardless of the number of medical complications. B) In the occipital segment of the corpus callosum (CC-Occ), medical complications moderated the association of FA and composite language score. A relation between CC-Occ FA and language at age 2 was found in the participants with 1 or 2 and with 3 or more medical risk factors.

3.6 Secondary Analyses

3.6.1 Along tract profile associations between language outcomes and white matter microstructural metrics

Results of secondary along-tract analyses of 2-year language outcomes and along-tract FA and MD are visualized in **Inline Supplementary Figure S1**. Results of the along-tract analysis additionally identified moderate correlations between FA and composite language scores in ILF-L nodes 9-15 that survived correction for multiple comparisons along the 30 nodes in the tract. The correlation between FA of the ILF-L nodes 9-15 and composite language score was 0.53 (95% CI = 0.17 – 0.77, p = 0.007).

3.6.2 Correlations of AD and RD with language

Composite language scores negatively correlated with mean AD of the CC-Occ (r = -0.418, 95% CI = -0.68 - -0.05, p = 0.02. The other associations between composite language and mean AD of the ILF-L, mean RD of the ILF-L, and mean RD of the CC-Occ were not significant.

4. DISCUSSION

In this investigation of 2-year-old children, we evaluated relations between white matter microstructure metrics from infant dMRI obtained prior to hospital discharge and later language outcomes, accounting for the moderating effect of number of neonatal medical complications. The major findings of this study were that language scores at age 2 were moderately positively correlated with mean FA of the ILF-L and moderately negatively correlated with mean FA of the CC-Occ. The mean tract associations were

not sufficiently strong to achieve significance after correction for multiple comparisons, but confidence intervals did not cross zero. Secondary along tract analyses reflected the mean tract correlation results and identified correlations with language in a region of the ILF-L. We identified a moderating effect of number of medical complications on the association between language and FA in CC-Occ, but not on the ILF-L.

The association between 2-year language outcomes and FA of the ILF-L is consistent with associations in older children born at term (Yeatman et al., 2012a) and preterm (Feldman et al., 2012; Murner-Lavanchy et al., 2018). The ventral pathways, including the ILF, are associated with mapping auditory speech sounds to meaning and simple syntactic structure processing (Hickok and Poeppel, 2007). The associations are consistent with reported deficits in receptive and expressive language skills in children born PT compared with their FT counterparts (Barre et al., 2011).

Language outcomes were negatively associated with FA of the CC-Occ, a white matter region prone to injury in children born preterm (Thompson et al., 2011). The correlation may indicate that FA in posterior periventricular regions indexes general white matter health or injury. Alternatively, or in addition, at these very young ages, the posterior corpus callosum may be relevant to language learning, as has been found later in childhood. The observed pattern for negative correlations within the posterior callosal pathway is consistent with reported negative associations between FA and readings skills in older PT children (Dubner et al., 2019) and individuals born FT (Dougherty et al., 2007; Frye et al., 2008; Hasan et al., 2012; Huber et al., 2019; Odegard et al., 2009). Diffusion metrics including FA and MD index multiple tissue properties including but not limited to myelin (Basser and Pierpaoli, 1996; Jones and

Cercignani, 2010). Thus, the observed white matter-language relations in the CC-Occ may also reflect variations in axonal properties, such as fiber densities, (Huber et al., 2019) fiber coherence or increased axonal diameter, which have been shown to occur in different proportions across regions of the corpus callosum (Aboitiz et al., 1992). A negative correlation between language and mean AD in the corpus callosum in this sample differs from previous work in 9-16 year old children born preterm in which associations between reading measures and FA were generally accompanied by associations between reading measures and RD, but not AD (Travis et al., 2016). White matter microstructure changes over time as children grow (Lebel and Deoni, 2018). FA may be indexing different axonal properties in infants than in older children. Future studies employing MRI techniques more specific to myelin (Mezer et al., 2013) and analytic approaches for interrogating axonal properties using dMRI data (Rokem et al., 2015; Zhang et al., 2012) are required to clarify the relations between tissue properties across regions of the corpus callosum and later outcomes in children born preterm. The combination of such methods and analytic approaches in future studies may also be important for interrogating the underlying tissue properties contributing the variations in FA and MD at these early stages. Such studies could also help to clarify why 2-year language outcomes were most related to variations in FA but not MD in our sample.

We found that FA of the ILF-L explained additional variance in language beyond number of medical complications. We speculate that the unique contribution of white matter microstructure and the absence of moderation by medical complications in this pathway may relate to the importance of the maturational state of oligodendrocyte precursors relative to the timing and location of oxidative-stress induced brain injuries.

For example, prolonged myelination of the ILF (Dubois et al., 2014) may reduce the susceptibility of such regions to injury in the neonatal period. Meanwhile the number of medical complications moderated the tract-language association in the posterior corpus callosum, with greater contribution of FA of the CC-Occ to prediction of language in children with more complications. The posterior corpus callosum myelinates earlier than the ILF (Dubois et al., 2014) and is in a region more prone to injury in the neonatal period. A child with more medical complications in the neonatal period may have greater injury to oligodendrocyte progenitors (Volpe, 2009) in the posterior periventricular region. Our findings suggest that PT children with fewer medical complications may rely primarily on language-related white matter pathways similar to FT children. For children with more medical complications, white matter may be more widely affected and therefore regions beyond traditional language pathways may contribute to variation in language development.

The study has several limitations. The sample of neonates whose data contributed to the analyses was modest. Using a stringent threshold for head motion, fewer than half of the scans from the original cohort could be used for analyses. The number of individual tracts included in the analyses was reduced further because of the technical challenges of segmenting the tracts using deterministic tractography. Given the final sample size, correlations between mean tract metrics and language did not reach statistical significance after correction for multiple comparisons. However, we prespecified the tracts that we planned to analyze based on the hypothesis that these specific tracts were likely to be associated with language measures. In addition, correlation coefficient confidence intervals did not cross zero and along tract analyses in

the ILF-L supported the mean tract findings. The results are therefore less likely to be due to chance alone. Future studies should enroll a large sample to increase confidence in the results.

We restricted analyses to tracts likely to be associated with language. We did not evaluate the whole brain to determine the specificity of our findings. Future studies should consider assessing the entire brain to confirm the findings described here and to explore whether other areas in the brain are related to language outcomes. We did not have information about other variables associated with language outcomes, such as socioeconomic status (SES), that may contribute to outcomes or may be confounded with other variables. Children with more medical complications may have been born to mothers with lower SES, which may be the true moderator of the relation between FA of the CC-Occ and language. Finally, we used a simplistic measure of neonatal illness. While conditions such as necrotizing enterocolitis, bronchopulmonary dysplasia, sepsis, and intraventricular hemorrhage contribute to hypoxia, ischemia, and inflammation, the exact mechanism by which these conditions lead to changes in white matter metrics is unknown. These major medical complications may directly affect white matter, or they may be a proxy for other disturbances such as prolonged poor nutritional status. Medical complications may also be a marker for other environmental influences associated with altered white matter microstructure, such as increased number of painful procedures (Vinall et al., 2014). These related endogenous or exogenous insults may be the ultimate culprits affecting white matter development.

5. CONCLUSION

In summary, a sensitive tractography method identified relations between language skills at age 2 years and infant dMRI white matter metrics from scans obtained prior to hospital discharge. Neonatal medical complications moderated the relations between white matter microstructure metrics in the posterior corpus callosum, suggesting that the white matter properties associated with language outcomes may relate to location and stage of myelination at injury and at the time of scan. Improvements in image acquisition and processing will potentially improve the clinical utility of this approach for identifying early markers of language development. White matter at near-term is one of many possible factors influencing later language, with the post-NICU environment having a large role (Adams et al., 2018). The detectable association between near-term white matter metrics and later language reported here requires further in-depth exploration of the contributors to language development and their interactions in larger samples of preterm infants.

Acknowledgments: We thank the children and families who participated in the study; K Yeom for assistance and expertise in reviewing clinical imaging; K Cunanan and K Vitale of the Stanford University Quantitative Sciences Unit for assistance with the statistical approach.

FUNDING: This work was supported by the National Institutes of Health [NICHD grant RO1-HD069162], and the Health Resources and Services Administration Maternal Child Health Bureau [T77MC09796] to Heidi M. Feldman, PI, the National Institutes of Health

[5K99HD084749] and the 2014 Society for Developmental and Behavioral Pediatrics Young Investigator Award to Katherine E. Travis, PI, and the Stanford Maternal Child Health Research Institute Tashia and John Morgridge Endowed Postdoctoral Fellowship to Sarah E. Dubner. This research was also supported by the Chiesi Foundation, Parma Italy; NIH Clinical and Translational Science Award UL1RR025744 for the Stanford Center for Clinical and Translational Education and Research (Spectrum) and for Stanford Center for Clinical Informatics, Stanford Translational Research Integrated Database Environment (STRIDE); Lucile Packard Foundation for Children's Health; NSF Graduate Research Fellowship grant [DGE-1147470], and by the Mary Baracchi Research Fund, Lucile Packard Children's Hospital at Stanford.

REFERENCES

Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E., 1992. Fiber composition of the human corpus callosum. Brain research 598, 143–153. https://doi.org/10.1016/0006-8993(92)90178-C

Adams, K.A., Marchman, V.A., Loi, E.C., Ashland, M.D., Fernald, A., Feldman, H.M., 2018. Caregiver Talk and Medical Risk as Predictors of Language Outcomes in Full Term and Preterm Toddlers. Child Dev 89, 1674–1690. https://doi.org/10.1111/cdev.12818

Bach, M., Laun, F.B., Leemans, A., Tax, C.M.W., Biessels, G.J., Stieltjes, B., Maier-Hein, K.H., 2014. Methodological considerations on tract-based spatial statistics (TBSS). NeuroImage 100, 358–369. https://doi.org/10.1016/j.neuroimage.2014.06.021

Back, S.A., 2015. Brain Injury in the Preterm Infant: New Horizons for Pathogenesis and Prevention. Pediatr Neurol 53, 185–92. https://doi.org/10.1016/j.pediatrneurol.2015.04.006

Ball, G., Boardman, J.P., Aljabar, P., Pandit, A., Arichi, T., Merchant, N., Rueckert, D., Edwards, A.D., Counsell, S.J., 2013. The influence of preterm birth on the developing thalamocortical connectome. Cortex 49, 1711–21. https://doi.org/10.1016/j.cortex.2012.07.006

Barre, N., Morgan, A., Doyle, L.W., Anderson, P.J., 2011. Language Abilities in Children Who Were Very Preterm and/or Very Low Birth Weight: A Meta-Analysis. The Journal of Pediatrics 158, 766-774.e1. https://doi.org/10.1016/j.jpeds.2010.10.032

Basser, P.J., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J Magn Reson B 111, 209–19.

Bayley, N., 2006. Bayley Scales of Infant and Toddler Development, 3rd ed. Harcourt Assessment Inc., San Antonio, TX.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289–300.

Bruckert, L., Shpanskaya, K., McKenna, E.S., Borchers, L.R., Yablonski, M., Blecher, T., Ben-Shachar, M., Travis, K.E., Feldman, H.M., Yeom, K.W., 2019. Age-Dependent White Matter Characteristics of the Cerebellar Peduncles from Infancy Through Adolescence. Cerebellum 18, 372–387. https://doi.org/10.1007/s12311-018-1003-9

Catani, M., Allin, M.P.G., Husain, M., Pugliese, L., Mesulam, M.M., Murray, R.M., Jones, D.K., 2007. Symmetries in human brain language pathways correlate with verbal recall. Proc. Natl. Acad. Sci. U.S.A. 104, 17163–17168. https://doi.org/10.1073/pnas.0702116104

Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: robust estimation of tensors by outlier rejection. Magn Reson Med 53, 1088–95. https://doi.org/10.1002/mrm.20426

Connor, C.M., Alberto, P.A., Compton, D.L., O'Connor, R.E., 2014. Improving Reading Outcomes for Students with or at Risk for Reading Disabilities: A Synthesis of the Contributions from the Institute of Education Sciences Research Centers. NCSER 2014-3000. National Center for Special Education Research.

Dodson, C.K., Travis, K.E., Ben-Shachar, M., Feldman, H.M., 2017. White matter microstructure of 6-year old children born preterm and full term. Neuroimage Clin 16, 268–275. https://doi.org/10.1016/j.nicl.2017.08.005

Dodson, C.K., Travis, K.E., Borchers, L.R., Marchman, V.A., Ben-Shachar, M., Feldman, H.M., 2018. White matter properties associated with pre-reading skills in 6-year-old children born preterm and at term. Developmental medicine and child neurology 60, 695–702. https://doi.org/10.1111/dmcn.13783

Dougherty, R.F., Ben-Shachar, M., Deutsch, G.K., Hernandez, A., Fox, G.R., Wandell, B.A., 2007. Temporal-callosal pathway diffusivity predicts phonological skills in children. Proceedings of the National Academy of Sciences of the United States of America 104, 8556–61. https://doi.org/10.1073/pnas.0608961104

Dubner, S.E., Dodson, C.K., Marchman, V.A., Ben-Shachar, M., Feldman, H.M., Travis, K.E., 2019. White matter microstructure and cognitive outcomes in relation to neonatal inflammation in 6-year-old children born preterm. NeuroImage: Clinical 101832. https://doi.org/10.1016/j.nicl.2019.101832

Dubois, J., Dehaene-Lambertz, G., Kulikova, S., Poupon, C., Huppi, P.S., Hertz-Pannier, L., 2014. The early development of brain white matter: a review of imaging studies in fetuses, newborns and infants. Neuroscience 276, 48–71. https://doi.org/10.1016/j.neuroscience.2013.12.044

Feldman, H.M., Lee, E.S., Yeatman, J.D., Yeom, K.W., 2012. Language and reading skills in school-aged children and adolescents born preterm are associated with white matter properties on diffusion tensor imaging. Neuropsychologia 50, 3348–62. https://doi.org/10.1016/j.neuropsychologia.2012.10.014

Feldman, H.M., Yeatman, J.D., Lee, E.S., Barde, L.H., Gaman-Bean, S., 2010. Diffusion tensor imaging: a review for pediatric researchers and clinicians. Journal of developmental and behavioral pediatrics□: JDBP 31, 346–56. https://doi.org/10.1097/DBP.0b013e3181dcaa8b

Field, A., 2013. Discovering Statistics Using IBM SPSS Statistics, 4th Edition, 4th edition. ed. SAGE Publications Ltd.

Foster-Cohen, S., Edgin, J.O., Champion, P.R., Woodward, L.J., 2007. Early delayed language development in very preterm infants: evidence from the MacArthur-Bates CDI. J Child Lang 34, 655–675.

Frye, R.E., Hasan, K., Xue, L., Strickland, D., Malmberg, B., Liederman, J., Papanicolaou, A., 2008. Splenium microstructure is related to two dimensions of reading skill. Neuroreport 19, 1627–1631. https://doi.org/10.1097/WNR.0b013e328314b8ee

Hasan, K.M., Molfese, D.L., Walimuni, I.S., Stuebing, K.K., Papanicolaou, A.C., Narayana, P.A., Fletcher, J.M., 2012. Diffusion tensor quantification and cognitive correlates of the macrostructure and microstructure of the corpus callosum in typically developing and dyslexic children. NMR Biomed 25, 1263–1270. https://doi.org/10.1002/nbm.2797

Hayes, A.F., 2017. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, 2nd ed. Guilford Press.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402. https://doi.org/10.1038/nrn2113

Huang, H., Zhang, J., Jiang, H., Wakana, S., Poetscher, L., Miller, M.I., van Zijl, P.C., Hillis, A.E., Wytik, R., Mori, S., 2005. DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. Neuroimage 26, 195–205. https://doi.org/10.1016/j.neuroimage.2005.01.019

Huber, E., Henriques, R.N., Owen, J.P., Rokem, A., Yeatman, J.D., 2019. Applying microstructural models to understand the role of white matter in cognitive development. Developmental Cognitive Neuroscience 36, 100624. https://doi.org/10.1016/j.dcn.2019.100624

Jones, D.K., Cercignani, M., 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR Biomed 23, 803–20. https://doi.org/10.1002/nbm.1543

Langer, N., Peysakhovich, B., Zuk, J., Drottar, M., Sliva, D.D., Smith, S., Becker, B.L., Grant, P.E., Gaab, N., 2017. White Matter Alterations in Infants at Risk for Developmental Dyslexia. Cerebral cortex (New York, N.Y.□: 1991) 27, 1027–1036. https://doi.org/10.1093/cercor/bhv281

Lebel, C., Beaulieu, C., 2009. Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. Human Brain Mapping 30, 3563–3573. https://doi.org/10.1002/hbm.20779

Lebel, C., Deoni, S., 2018. The development of brain white matter microstructure. Neuroimage 182, 207–218. https://doi.org/10.1016/j.neuroimage.2017.12.097

Leemans, A., Jones, D.K., 2009. The B-matrix must be rotated when correcting for subject motion in DTI data. Magn Reson Med 61, 1336–1349. https://doi.org/10.1002/mrm.21890

Marchman, V.A., Adams, K.A., Loi, E.C., Fernald, A., Feldman, H.M., 2016. Early language processing efficiency predicts later receptive vocabulary outcomes in children born preterm. Child Neuropsychol 22, 649–65. https://doi.org/10.1080/09297049.2015.1038987

Marchman, V.A., Ashland, M.D., Loi, E.C., Adams, K.A., Fernald, A., Feldman, H.M., 2019. Predictors of early vocabulary growth in children born preterm and full term: A study of processing speed and medical complications. Child Neuropsychol 1–21. https://doi.org/10.1080/09297049.2019.1569608

Mezer, A., Yeatman, J.D., Stikov, N., Kay, K.N., Cho, N.J., Dougherty, R.F., Perry, M.L., Parvizi, J., Hua le, H., Butts-Pauly, K., Wandell, B.A., 2013. Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. Nat Med 19, 1667–72. https://doi.org/10.1038/nm.3390

Mishra, A., Anderson, A.W., Wu, X., Gore, J.C., Ding, Z., 2010. An improved Bayesian tensor regularization and sampling algorithm to track neuronal fiber pathways in the language circuit. Med Phys 37, 4274–4287. https://doi.org/10.1118/1.3456113

Murner-Lavanchy, I.M., Kelly, C.E., Reidy, N., Doyle, L.W., Lee, K.J., Inder, T., Thompson, D.K., Morgan, A.T., Anderson, P.J., 2018. White matter microstructure is associated with language in children born very preterm. Neuroimage Clin 20, 808–822. https://doi.org/10.1016/j.nicl.2018.09.020

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Human Brain Mapping 15, 1–25. https://doi.org/10.1002/hbm.1058

Odegard, T.N., Farris, E.A., Ring, J., McColl, R., Black, J., 2009. Brain connectivity in non-reading impaired children and children diagnosed with developmental dyslexia. Neuropsychologia 47, 1972–7. https://doi.org/10.1016/j.neuropsychologia.2009.03.009

Oouchi, H., Yamada, K., Sakai, K., Kizu, O., Kubota, T., Ito, H., Nishimura, T., 2007. Diffusion anisotropy measurement of brain white matter is affected by voxel size: underestimation occurs in areas with crossing fibers. AJNR Am J Neuroradiol 28, 1102–1106. https://doi.org/10.3174/ajnr.A0488

Perani, D., Saccuman, M.C., Scifo, P., Anwander, A., Awander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., Friederici, A.D., 2011. Neural language networks at birth. Proc. Natl. Acad. Sci. U.S.A. 108, 16056–16061. https://doi.org/10.1073/pnas.1102991108

Rohde, G.K., Barnett, A.S., Basser, P.J., Marenco, S., Pierpaoli, C., 2004. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. Magn Reson Med 51, 103–14. https://doi.org/10.1002/mrm.10677

Rokem, A., Yeatman, J.D., Pestilli, F., Kay, K.N., Mezer, A.A., van der Walt, S., Wandell, B.A., 2015. Evaluating the Accuracy of Diffusion MRI Models in White Matter. PLOS ONE 10, e0123272. https://doi.org/10.1371/journal.pone.0123272

Rose, J., Cahill-Rowley, K., Vassar, R., Yeom, K.W., Stecher, X., Stevenson, D.K., Hintz, S.R., Barnea-Goraly, N., 2015. Neonatal brain microstructure correlates of

neurodevelopment and gait in preterm children 18-22 mo of age: an MRI and DTI study. Pediatric research 78, 700–8. https://doi.org/10.1038/pr.2015.157
Salvan, P., Tournier, J.D., Batalle, D., Falconer, S., Chew, A., Kennea, N., Aljabar, P., Dehaene-Lambertz, G., Arichi, T., Edwards, A.D., Counsell, S.J., 2017. Language ability in preterm children is associated with arcuate fasciculi microstructure at term. Hum Brain Mapp 38, 3836–3847. https://doi.org/10.1002/hbm.23632
Sanders, L.M., Federico, S., Klass, P., Abrams, M.A., Dreyer, B., 2009. Literacy and Child Health: A Systematic Review. Arch Pediatr Adolesc Med 163, 131–140. https://doi.org/10.1001/archpediatrics.2008.539
Shim, S.Y., Jeong, H.J., Son, D.W., Jeong, J.S., Oh, S.H., Park, S.Y., Ryu, T.H., Kim, Y.B., Cho, Z.H., 2012. Altered Microstructure of White Matter Except the Corpus Callosum Is Independent of Prematurity. Neonatology 102, 309–315.
Singer, L.T., Siegel, A.C., Lewis, B., Hawkins, S., Yamashita, T., Baley, J., 2001. Preschool language outcomes of children with history of bronchopulmonary dysplasia and very low birth weight. J Dev Behav Pediatr 22, 19–26.
Thompson, D.K., Inder, T.E., Faggian, N., Johnston, L., Warfield, S.K., Anderson, P.J., Doyle, L.W., Egan, G.F., 2011. Characterization of the corpus callosum in very preterm and full-term infants utilizing MRI. Neuroimage 55, 479–90. https://doi.org/10.1016/j.neuroimage.2010.12.025
Travis, K.E., Adams, J.N., Ben-Shachar, M., Feldman, H.M., 2015a. Decreased and Increased Anisotropy along Major Cerebral White Matter Tracts in Preterm Children and Adolescents. PloS one 10, e0142860. https://doi.org/10.1371/journal.pone.0142860
Travis, K.E., Adams, J.N., Kovachy, V.N., Ben-Shachar, M., Feldman, H.M., 2017. White matter properties differ in 6-year old Readers and Pre-readers. Brain Struct Funct 222, 1685–1703. https://doi.org/10.1007/s00429-016-1302-1
Travis, K.E., Ben-Shachar, M., Myall, N.J., Feldman, H.M., 2016. Variations in the neurobiology of reading in children and adolescents born full term and preterm. Neuroimage Clin 11, 555–65. https://doi.org/10.1016/j.nicl.2016.04.003
Travis, K.E., Golden, N.H., Feldman, H.M., Solomon, M., Nguyen, J., Mezer, A., Yeatman, J.D., Dougherty, R.F., 2015b. Abnormal white matter properties in adolescent girls with anorexia nervosa. Neuroimage Clin 9, 648–59. https://doi.org/10.1016/j.nicl.2015.10.008
van Noort-van der Spek, I.L., Franken, M.-C.J.P., Weisglas-Kuperus, N., 2012. Language Functions in Preterm-Born Children: A Systematic Review and Meta-analysis. Pediatrics 129, 745–754. https://doi.org/10.1542/peds.2011-1728
Vinall, J., Miller, S.P., Bjornson, B.H., Fitzpatrick, K.P.V., Poskitt, K.J., Brant, R., Synnes, A.R., Cepeda, I.L., Grunau, R.E., 2014. Invasive procedures in preterm children: brain and cognitive development at school age. Pediatrics 133, 412–421. https://doi.org/10.1542/peds.2013-1863
Volpe, J.J., 2009. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol 8, 110–24. https://doi.org/10.1016/S1474-4422(08)70294-1

Wocadlo, C., Rieger, I., 2007. Phonology, rapid naming and academic achievement in very preterm children at eight years of age. Early Human Development 83, 367–377. https://doi.org/10.1016/j.earlhumdev.2006.08.001

Yeatman, J.D., Dougherty, R.F., Ben-Shachar, M., Wandell, B.A., 2012a. Development of white matter and reading skills. Proceedings of the National Academy of Sciences of the United States of America 109, E3045-53. https://doi.org/10.1073/pnas.1206792109

Yeatman, J.D., Dougherty, R.F., Myall, N.J., Wandell, B.A., Feldman, H.M., 2012b. Tract profiles of white matter properties: automating fiber-tract quantification. PloS one 7, e49790. https://doi.org/10.1371/journal.pone.0049790

Yeatman, J.D., Dougherty, R.F., Rykhlevskaia, E., Sherbondy, A.J., Deutsch, G.K., Wandell, B.A., Ben-Shachar, M., 2011. Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. Journal of cognitive neuroscience 23, 3304–3317.

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., Fischl, B., 2014. Spurious group differences due to head motion in a diffusion MRI study. NeuroImage 88, 79–90. https://doi.org/10.1016/j.neuroimage.2013.11.027

Young, J.M., Morgan, B.R., Whyte, H.E.A., Lee, W., Smith, M.L., Raybaud, C., Shroff, M.M., Sled, J.G., Taylor, M.J., 2017. Longitudinal Study of White Matter Development and Outcomes in Children Born Very Preterm. Cereb. Cortex 27, 4094–4105. https://doi.org/10.1093/cercor/bhw221

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 61, 1000–16. https://doi.org/10.1016/j.neuroimage.2012.03.072