NeuroImage 192 (2019) 145–155

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

White matter connectomes at birth accurately predict cognitive abilities at age 2

Jessica B. Girault [a], Brent C. Munsell [b], Danaële Puechmaille [a], Barbara D. Goldman [c], Juan C. Prieto [a], Martin Styner [a], John H. Gilmore [a,*]

[a] Department of Psychiatry, UNC Chapel Hill, Chapel Hill, NC, 27599, USA
[b] Department of Computer Science, College of Charleston, Charleston, SC, 29424, USA
[c] Department of Psychology & Neuroscience, UNC Chapel Hill, Chapel Hill, NC, 27599, USA

ARTICLE INFO

Keywords:
White matter
Connectome
Infant brain
Cognition
Prediction
Biomarker

ABSTRACT

Cognitive ability is an important predictor of mental health outcomes that is influenced by neurodevelopment. Evidence suggests that the foundational wiring of the human brain is in place by birth, and that the white matter (WM) connectome supports developing brain function. It is unknown, however, how the WM connectome at birth supports emergent cognition. In this study, a deep learning model was trained using cross-validation to classify full-term infants ($n = 75$) as scoring above or below the median at age 2 using WM connectomes generated from diffusion weighted magnetic resonance images at birth. Results from this model were used to predict individual cognitive scores. We additionally identified WM connections important for classification. The model was also evaluated in a separate set of preterm infants ($n = 37$) scanned at term-age equivalent. Findings revealed that WM connectomes at birth predicted 2-year cognitive score group with high accuracy in both full-term (89.5%) and preterm (83.8%) infants. Scores predicted by the model were strongly correlated with actual scores ($r = 0.98$ for full-term and $r = 0.96$ for preterm). Connections within the frontal lobe, and between the frontal lobe and other brain areas were found to be important for classification. This work suggests that WM connectomes at birth can accurately predict a child's 2-year cognitive group and individual score in full-term and preterm infants. The WM connectome at birth appears to be a useful neuroimaging biomarker of subsequent cognitive development that deserves further study.

1. Introduction

Individual differences in cognitive ability in youth have been associated with risk for poor mental health outcomes in adulthood, with lower performing children at increased risk for developing disorders including schizophrenia (Dickson et al., 2012; Gunnell et al., 2002), post-traumatic stress disorder (Gale et al., 2008; Koenen et al., 2007), and depression (Gale et al., 2008; Zammit et al., 2004). Cognitive scores in toddlerhood are correlated with school-age ability (Bayley, 1949; Bishop et al., 2003; Girault et al., 2018b; McCall et al., 1972; Stumm et al., 2009), which is a fairly stable marker of intelligence in adulthood (Bradway and C. W. Thompson, 1962; Deary et al., 2013, 2004; McCall, 1977), suggesting the foundations of individual differences in ability are likely determined in the first years of life. However, relatively little is known about how the brain develops to support emergent cognition, or how useful neuroimaging biomarkers of early brain organization may be for predicting future cognitive abilities.

WM connections play an important role in cognition (Zatorre et al., 2012) by optimizing rapid information transfer between brain areas (Mabbott et al., 2006; Nagy et al., 2004). At birth, white matter (WM) in the human brain is a highly connected network of largely unmyelinated axons that will serve as the foundation upon which future fine-tuning of cortical circuitry takes place via processes that include synaptogenesis, dendritic arborization, and myelination (Dubois et al., 2014). By week 30 of gestation, major pathways underlying the rich-club organization of the brain are established (Ball et al., 2014), and by birth WM networks exhibit a small world architecture (Yap et al., 2011). This suggests that the foundational wiring of brain circuitry is established *in utero* and is largely in place by the time of term birth, a finding which has been supported by tractography studies (Dubois et al., 2008; Huang et al., 2006). The WM connectome, as a physical network, has important implications for both cortical structural development (Essen, 1997) and

* Corresponding author.University of North Carolina at Chapel Hill, Campus Box # 7160, Chapel Hill, NC, 27599-7160, USA.
E-mail address: john_gilmore@med.unc.edu (J.H. Gilmore).

https://doi.org/10.1016/j.neuroimage.2019.02.060
Received 7 October 2018; Received in revised form 18 February 2019; Accepted 22 February 2019
Available online 27 February 2019

functional brain connectivity (Hagmann et al., 2010; Park and Friston, 2013; Sporns, 2013).

The WM connectome is more adult-like at birth than the functional connectome, with widespread structural hubs already present in medial frontal, parietal, and hippocampal areas (Huang et al., 2015; van den Heuvel et al., 2015) along with regions in the posterior cingulate and insula (Ball et al., 2014). In contrast, functional networks at birth are comprised of a relatively immature network of hubs in primary sensory, auditory, and sensorimotor areas (Cao et al., 2017a; Fransson et al., 2010). Interestingly, cross-sectional developmental studies have shown that coupling between WM and functional networks increases from 30 weeks gestation into adulthood (Hagmann et al., 2010; van den Heuvel et al., 2015). This body of work highlights the possibility that early-maturing WM connectomes serve as the initial foundation upon which diverse functional networks are built (Cao et al., 2017b).

Importantly, the WM connectome must provide enough flexibility to support dynamic large-scale functional reorganization that has been shown to occur during cognitive tasks (Cohen and D'Esposito, 2016). Recent studies have begun to reveal interesting links between individual differences in WM connectomic features in pediatric populations and future cognition and behavior. WM connections between the thalamus and cortex have been related to cognitive abilities at age 2 in preterm infants (Ball et al., 2015), and WM connectomes at birth have been used to derive subject communities that were related to maternal reports of child behaviors at ages 2 and 4 in infants born full-term (Wee et al., 2016). A recently developed methodological approach outlined the utility of deep convolutional neural networks in predicting cognitive and motor scores at age 2 from WM connectomes at birth in very preterm infants (Kawahara et al., 2017). This recent work suggests that WM connectivity at birth is important for future cognitive and behavioral outcomes in toddlerhood, though to date, predictive relationships have been poor to modest and have been conducted in full-term or preterm-only samples.

In the present study, we extend our previous work investigating the associations between WM tractography and cognition, which found that heritable components of white matter are associated with cognition in early life (Lee et al., 2017) and that these brain-cognition associations are of a global, non-specific nature and detectable as early as birth (Girault et al., 2018a). Together, these studies found that WM integrity from any single fiber tract, or white matter component, at birth carried relatively little information about future cognitive abilities, though based on the global nature of the findings we hypothesized that WM may be a useful biomarker if it were considered as a global system, using a higher-dimensional analytic approach. Further, using machine learning to predict future cognitive outcomes from neuroimaging markers taken shortly after birth is highly clinically-relevant, and part of an emerging body of work predicting outcomes in high-risk infant populations (Emerson et al., 2017; Hazlett et al., 2017a; Kawahara et al., 2017). Here we extend this framework to determine if WM connectomes at birth can be used to predict individual differences in cognitive abilities at age 2, across a period of rapid, dynamic brain development (Geng et al., 2012; Gilmore et al., 2018; Knickmeyer et al., 2008; Lyall et al., 2015), in a heterogeneous sample of infants followed longitudinally.

We used a deep learning approach to determine the predictive ability of WM connectomes at birth for subsequent cognition and identify features of the WM connectome at birth important for classification accuracy. Importantly, we tested the robustness of our prediction approach by training with full-term born infants and applying the resulting model to an independent set of preterm infants scanned at approximate term-age-equivalent. If our model trained on data from full-term infants could be applied to a sample of infants born prematurely and achieve similar accuracy, it would suggest that there exists an underlying set of WM connectomic features that lay the foundation for emergent cognition across late gestation into the neonatal period.

2. Methods and materials

2.1. Sample

Participants were part of the UNC Early Brain Development Study (Gilmore et al., 2010; Knickmeyer et al., 2008, 2016). We retrospectively identified 115 infants (twins and singletons) with high-quality diffusion and structural MRIs at birth and cognitive assessments at age 2. This dataset contained infants born between 212 and 295 days (30.28–42.14 weeks) gestation, as mothers were recruited prenatally and the gestational age of the infant was not exclusionary. To avoid introducing age variability in the training and testing datasets, primary model building and classification were performed using full-term (FT) infants (≥37 weeks gestation; $n = 75$) and models were then applied to preterm (PT) infants (<37 weeks; $n = 37$). Applying the model trained on FT data to the unique PT infants allowed us to achieve our goals of (1) validating this model in an independent dataset and (2) testing its robustness to classifying unseen connectome datasets with features that may vary based on gestational age, and thus are potentially more heterogeneous than those used for training. However, it is important to note that all infants were scanned shortly after birth, or as close to term-age-equivalent as possible; the youngest infant was scanned at a gestational age of 37.4 weeks, which was 50 days following their preterm birth at 30.28 gestational weeks.

Demographics are presented in Table 1; gestational age distributions are presented in Fig. 1A. Informed written consent was obtained from a parent or legal guardian of each participant. This study was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill. The PT subject sample included infants with perinatal complications including stay in the neonatal intensive care unit (PT: $n = 23$) and failure to thrive (PT: $n = 1$). Additionally, one preterm infant

Table 1
Participant demographic information.

	Full Term (N = 75)	Preterm (N = 37)	*T-test*[a]	*df*	*P-value*
	Mean (SD)	*Mean (SD)*			
Gestational Age Birth (days)	272.72 (9.27)	238.68 (14.21)	13.25	51.60	<0.0001
Birth Weight (grams)	3101.24 (475.31)	2188.22 (485.13)	9.43	70.48	<0.0001
Stay in NICU (days)	0 (0)	10.89 (12.37)	−5.35	36	<0.0001
Chronological Age at MRI (days)	22.72 (9.44)	45.19 (14.47)	−8.59	51.63	<0.0001
Gestational Age at MRI (days)	295.16 (9.87)	283.11(10.97)	5.65	65.46	<0.0001
Age at 2yr MSEL (days)	747.53 (22.41)	779.08 (25.14)	−6.47	64.91	<0.001
2yr ELC	107.11 (15.86)	105.19 (15.85)	0.60	71.84	0.549
Maternal Age (years)	30.02 (5.76)	29.95 (6.14)	0.07	67.78	0.947
Paternal Age (years)	32.49 (6.00)	32.0 (8.40)	0.31	54.72	0.751
Maternal Education (years)	15.06 (3.25)	15.89 (3.23)	−1.28	72.24	0.205
Paternal Education (years)	15.06 (3.79)	15.36 (2.98)	−0.45	86.62	0.652
Household Income ($)	$73,094 (63,538)	$88,537 (64,870)	−1.16	66.53	0.250
	N	*N*	*Chi Sq.*		*P-value*
Male/Female	39/36	24/13	1.18		0.276
Singleton/Twin	47/28	4/33	24.81		<0.0001

[a] Calculated using Welch's two-sample, two-tailed *t*-test assuming unequal variance.

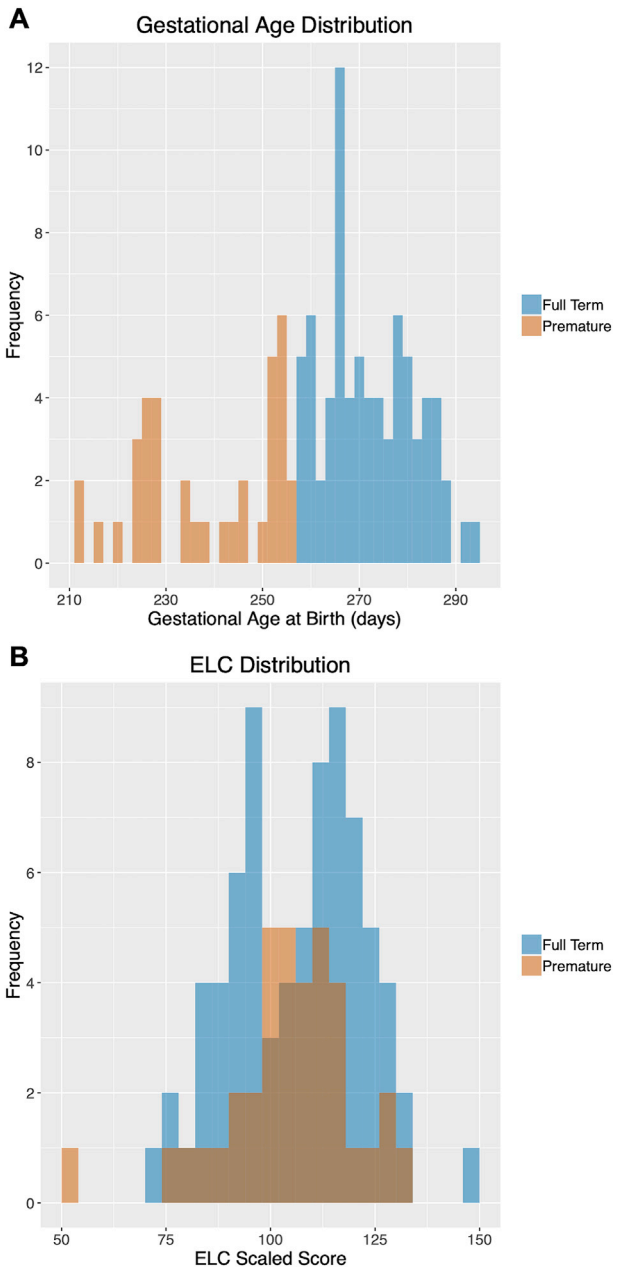

Fig. 1. Participant Gestational Age at Birth and ELC Distributions. Participant distributions for gestational age at birth (A) and ELC scores at age 2 (B) are presented. Data from full-term infants are shown in blue and preterm infants are shown in orange.

was identified as having a potential developmental delay due to scoring ≤70 on the 2-year cognitive assessment. FT subjects used for training the algorithm scored within a normal range of developmental assessments at age 2 and had no documented perinatal complications.

2.2. Cognitive assessment

Cognitive ability was assessed at age 2 using the Mullen Scales of Early Learning (Mullen, 1995). Measures of fine motor, visual reception, expressive and receptive language were collected by experienced testers. Age-standardized T-scores from these scales were combined into an Early Learning Composite (ELC) standardized score (range: 49–155, mean = 100, SD = 15). The ELC has high internal consistency (median = 0.91) and reliability (median = 0.84 for the cognitive scales during this age window), and principal factor loadings of the scales lend support for the construct validity of the ELC as a general measure of cognitive ability (Mullen, 1995). The range of ELC scores for FT and PT infants is shown in Fig. 1B.

2.3. Image acquisition

All images were acquired between 2009 and 2012 using a Siemens Allegra head-only 3T scanner (FT: $n = 55$, PT: $n = 27$) or a TIM Trio 3T scanner (FT: $n = 20$, PT: $n = 10$), which replaced the Allegra in 2011 (Siemens Medical System, Inc., Erlangen, Germany). There is no significant difference in the distribution of participants across scanners between the FT and PT group (X^2 (1, N = 112) = 3.10e^{-31}, p > 0.99). Infants were scanned during natural sleep, fitted with earplugs and secured using a vacuum-fixed immobilization device. Diffusion weighted images (DWIs) were acquired using a single-shot echo-planar imaging spin-echo sequence with an acquisition time of 6 min and 14 s. For all DWI data, 42 directions of diffusion sensitization were acquired with a b value of 1000 s/mm^2 in addition to seven baseline (b value = 0) images (generating a total of 49 DWIs). Acquisition parameters: TR/TR/Flip angle = 7680/82/90°, slice thickness = 2 mm, and in-plane resolution = 2 × 2 mm^2, with a total of 60–72 slices.

T1 and T2 weighted images were additionally acquired for the white matter surface generation and its propagation to DWI space. The T2-weighted images were acquired on the Allegra using a turbo-spin echo sequence (TSE, TR = 6200 ms, TE1 = 20 ms, TE2 = 119 ms, flip angle = 150°, spatial resolution = 1.25 mm × 1.25 mm x 1.95 mm; Type1: FT: $n = 5$, PT: $n = 1$) or a "fast" turbo-spin echo sequence was collected on the Allegra using a decreased TR, a smaller image matrix, and fewer slices (TSE, TR range = 5270 ms-5690 ms, TE1 range = 20 ms-21 ms, TE2 range = 119 ms-124 ms, flip angle = 150°, spatial resolution = 1.25 mm × 1.25 mm x 1.95 mm; Type2: FT: $n = 50$, PT: $n = 26$). For the Trio, participants were initially scanned using a TSE protocol (TR = 6200 ms, TE1 = 17, TE2 = 116 ms, flip angle = 150°, spatial resolution = 1.25 mm × 1.25 mm x 1.95 mm; Type3: FT: $n = 2$, PT: $n = 2$) while the rest were scanned using a 3DT2 SPACE protocol (TR = 3200 ms, TE = 406, flip angle = 120°, spatial resolution = 1 mm × 1 mm x 1 mm; Type 4: FT: $n = 18$, PT: $n = 8$). There is no significant difference in the distribution of FT and PT participants across T2 scan types used to generate WM surfaces (X^2 (3, N = 112) = 1.35, p = 0.716).

The T1-weighted images were acquired on the Allegra using a 3D magnetization prepared rapid gradient echo sequence (MP-RAGE TR = 1820 ms, TE = 4.38 ms, flip angle = 7°, spatial resolution = 1 mm × 1 mm x 1 mm, with matrix dimensions of 256 × 192 or 256 × 144). T1 images on the Trio were collected using a lower echo time (MP-RAGE TR = 1820 ms, TE = 3.75 ms, flip angle = 7°, spatial resolution = 1 mm × 1 mm x 1 mm).

While there was no significant difference between the distribution of FT and PT participants across the two scanners, we conducted a series of tests to ensure that using DWI data from two scanners (Allegra vs. Trio) did not influence our results. We tested whether ELC scores, or the mean absolute error (MAE) between the actual and predicted ELC scores differed between subjects scanned on either scanner ("scanner groups") and they did not (ELC scores, full sample: $t = 1.44$, $df = 59.66$, $p = 0.155$). Scanner groups also did not significantly differ in terms gestational age ($t = -0.361$, $df = 59.98$, $p = 0.720$), birth weight ($t = 0.285$, $df = 63.04$, $p = 0.778$), duration in the NICU ($t = 0.640$, $df = 57.24$, $p = 0.524$), gestational age at scan ($t = 0.658$, $df = 50.90$, $p = 0.514$), age at ELC testing ($t = -1.71$, $df = 44.53$, $p = 0.094$), parental age (maternal: $t = 0.134$, $df = 46.02$, $p = 0.892$; paternal: $t = -0.033$, $df = 50.83$, $p = 0.973$), parental education level (maternal: $t = -0.426$, $df = 46.12$, $p = 0.672$; paternal: $t = -0.482$, $df = 54{,}06$, $p = 0.632$), nor in the distribution gestation number (singleton vs. twin, (X^2 (1, N = 112) = 0.857, p = 0.354)). There was a marginally significant difference in the distribution of males and females across the scanners (X^2 (1, N = 112) = 3.96,

$p = 0.047$); where the ratio of males to females is equal on the Allegra (41 males, 41 females), but more males were scanned on the Trio (22 males, 8 females). Additionally, we tested whether there were differences in the WM connectomes generated from the two scanners by testing for a group effect of scanner across the connectome. After adjusting for sex, no significant effect of scanner was present (See Supplement Section I).

2.4. WM surface generation and propagation

Cortical surfaces were generated from the neonatal T2-weighted images as previously described (Jha et al., 2018; Li et al., 2016). An infant-specific pipeline (Li et al., 2013) was used to perform tissue segmentation and a deformable surface method (Li et al., 2014, 2012) was applied to reconstruct WM surfaces that were parcellated into 78 regions corresponding to those in the automated anatomical labeling (AAL) atlas adapted for the infant brain (Gilmore et al., 2012; Tzourio-Mazoyer et al., 2002). All surfaces were examined for accurate mapping. T1-weighted images were then rigidly co-aligned with the processed T2-weighted images. T2 and T1-weighted images were then jointly used to propagate the WM surfaces into DWI space via deformable symmetric diffeomorphic co-registration to the average baseline ($b = 0$) image and the axial diffusivity map via ANTS (Avants et al., 2011).

2.5. WM connectome generation

A study-specific quality control protocol was applied to all DWI data as outlined in Supplement Section I. Probabilistic tractography was performed with CIVILITY (Puechmaille et al., 2017), which utilizes FSL tools bedpost and probtrackx2 (Behrens et al., 2007) to determine the diffusion connectome. Prior to tractography, Bayesian estimation of diffusion parameters was computed to allow for data-driven selection of the number of supported fiber orientations at each voxel, accounting for multiple orientations and crossing fibers (Behrens et al., 2007). For our analyses, we used two fiber orientations for voxel fitting. Each region from the WM surface was used as a seed region, and probabilistic tractography was performed using probtrackx2 (Behrens et al., 2007), with the number of streamlines per voxel set to 3000, a step-length of 0.75 mm, and seed sphere sampling size of 0.5 mm.

A 78×78 connectivity matrix (or connectome) was calculated using the probabilistic tractography data. Connectivity is defined as the number of probabilistic WM fiber tract streamlines arriving at region *j* when region *i* was seeded, averaged with the number of probabilistic WM fiber tract streamlines arriving at region *i* where region *j* was seeded. This step is iteratively repeated to ensure all 78 regions are treated as seed regions producing a connectivity matrix that is symmetric with respect to the main diagonal. Lastly, connectivity matrices were then normalized using a maximum scaling technique by dividing the entire matrix by its largest connectivity value. These symmetric, normalized matrices were used in the deep learning approach described below.

2.6. Deep learning pipeline overview

As illustrated in Fig. 2, our approach begins with preprocessing the infant WM connectomes to reshape them into connectivity feature vectors and generating classification labels to determine ELC 2-year score groups (above the median (AM) and below the median (BM)). There are two main steps in the prediction pipeline that follows: (1) connectivity feature vectors are input into the ELC classification model to identify median group membership (AM or BM) and then (2) the AM or BM classification probability value is directed to the corresponding 2-year ELC AM or BM prediction model and a 2-year ELC score is found. The rationale for this two-step approach and methodological details for the classification, score prediction, application to the preterm dataset, and generation of a connectivity fingerprint are outlined below.

2.7. Rationale for the 2-step approach

As illustrated in Fig. 2, our two-step approach includes the construction of two models: *classification* followed by *prediction*. In general, the two-step approach has been successfully incorporated into several machine learning designs (Choi and Jin, 2016; Hazlett et al., 2017b; Kim and B. Lee, 2018; Kooi et al., 2017; Suk et al., 2015) to mitigate overfitting and improve model generalization performance. Furthermore, we chose to use this two-step approach because preliminary experimentation showed that using a single neural network to predict ELC scores directly produces very poor results (see Supplement Section VI for details). Due to the small sample size and the large complexity of a direct prediction of the continuous ELC score, the results of the single neural network model were consistently and strongly biased towards a singular value. Based on such evidence of overfitting, a two-step pipeline was implemented to improve prediction performance.

The two-step pipeline design starts with an easier discrete *two-group* classification problem (AM or BM) that is then followed by a much more difficult *continuous score* prediction problem in which the classification accuracy associated with the probability that an infant will belong to the AM or BM group is directed to a linear regression model predicting the continuous ELC score. To better understand this design decision, the box plots in Fig. 3A show the distribution of classification probability values for infants assigned by the model to the *correct* median score group. The standard deviation of the classification probability is within ± 0.013 of the AM or BM group mean, suggesting the variability of the underlying data distribution is very small, and thus may not provide adequate information resolution to train a single prediction model. To resolve this, we take the classification probabilities shown by the red line in Fig. 3B from the first step in the pipeline and normalize them to a value in [0 1], shown by the green line in Fig. 3B before the ELC 2-year score linear regression prediction model is trained in the second step, thus transforming a large learning step into a smaller learning step.

2.8. ELC 2-year median score group classification model

The upper triangular portion of the 78×78 FT infant WM connectivity matrices were reshaped into one-dimensional connectivity feature vectors each containing 3003 WM connections. The connectivity feature vectors, along with their associated median score group ELC classification labels, were then used to train a dense neural network. Training was performed using a 10-fold cross-validation approach where the 78 FT infants were randomly divided into 10 folds (where the number of subjects assigned to each fold ranged from 6 to 9), maintaining the same ratio of outcomes (AM and BM). In an iterative fashion, nine folds are used to train the model and the one left-out fold is used test the model, this is continued until all combinations of folds have been used for training and testing, yielding 10 different prediction models (see Supplement Section III for further details).

Each neural network defines one input layer, five hidden layers, five activation layers, three dropout layers, and one output layer (see Supplemental Fig. 2). The input layer has 3003 neural network nodes, one for each WM connection in the connectivity feature vector, and the output layer has two neural network nodes, one for each ELC median score group. One additional supervised learning layer with two neural network nodes was added for the model training. Once the supervised training step was completed, the supervised training layer was removed, and the response value of neural network nodes in the output layer were used for ELC median score group classification and classification probability. For example, given a connectivity feature vector, if the output responses of the neural network are $AM = 0.35$ and $BM = 0.65$, then the infant is classified as BM with probability equal to 0.65. We calculated an overall accuracy value which quantified the percentage of times an infant was correctly assigned to their cognitive group averaged over the 10 trained models. Detailed information on the classification model is presented in Supplement Section III.

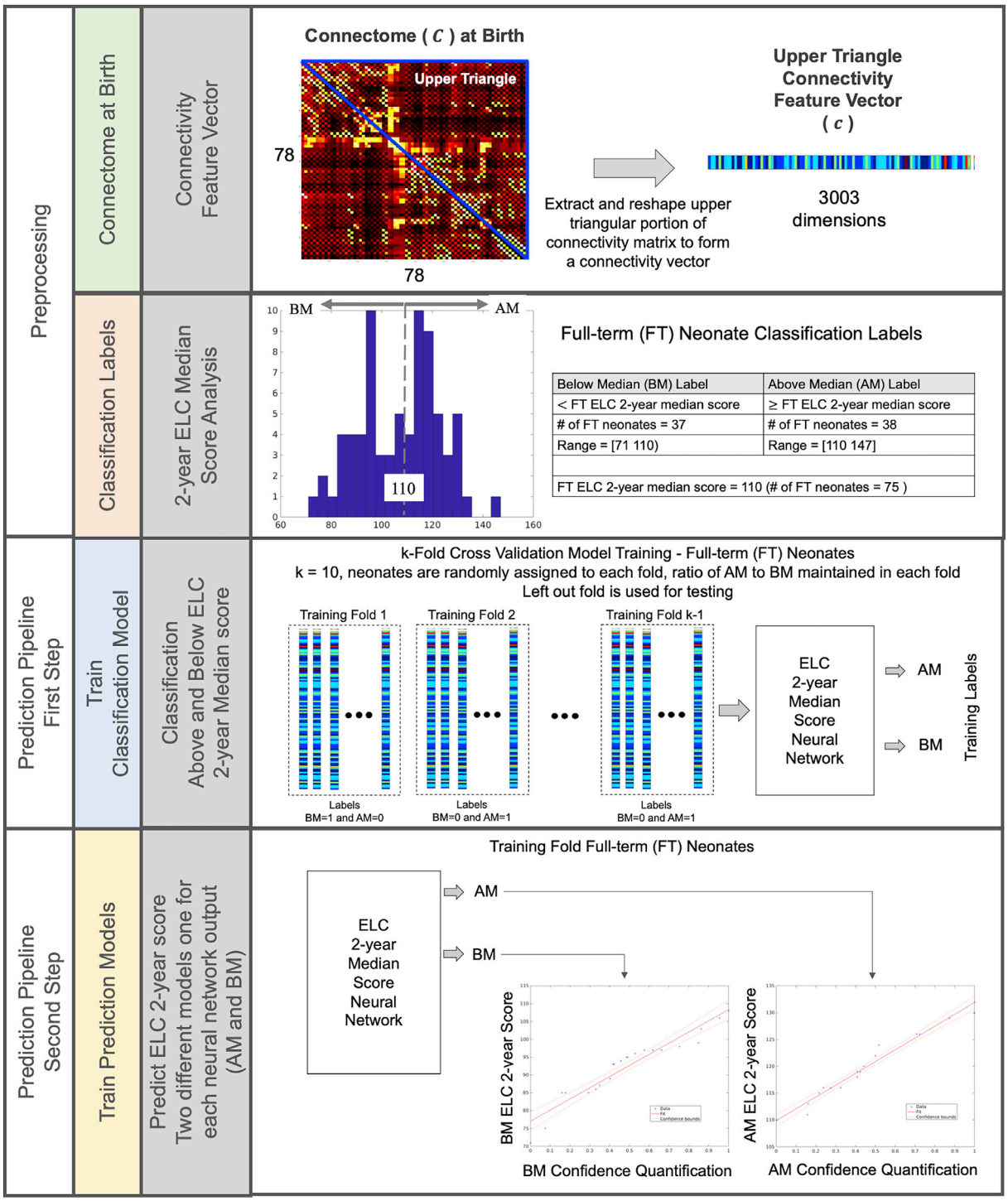

Fig. 2. Approach for predicting ELC scores at age 2 from WM connectomes at birth. Data were preprocessed to reshape each WM connectivity matrix at birth into a 3003 dimension connectivity feature vector ($(78 \times 77)/2 = 3003$ region pairs) using the normalized connectivity values in the upper triangular portion of the symmetric connectome (not including the main diagonal). Classification labels were generated by grouping 75 FT subjects into two categories based on their performance on the 2-year ELC relative to others in our sample: above the median (AM) and below the median (BM). Next, a 10-fold cross-validation approach was used to train and test a two-step prediction pipeline. First, a classification model was trained to identify ELC group (AM, BM), generating classification probability values. Classification probability values were then used to train two separate prediction models for each ELC median score group that directly estimates the ELC score at age 2.

2.9. ELC 2-year score prediction models

The classification probability values generated by the ELC 2-year median score group classification model were first normalized (Fig. 3) and then used to train two different linear regression models (one for the BM group and AM group) capable of predicting the actual ELC score at age 2. For each prediction model, one predictor variable (normalized classification probability) and one response variable (actual 2-year ELC score) were used to train the linear regression model. To test the strength of the regression prediction model, we calculated Pearson's product-moment correlations between infant's actual and predicted scores and tested for significance at the level of $p < 0.05$. We additionally calculated

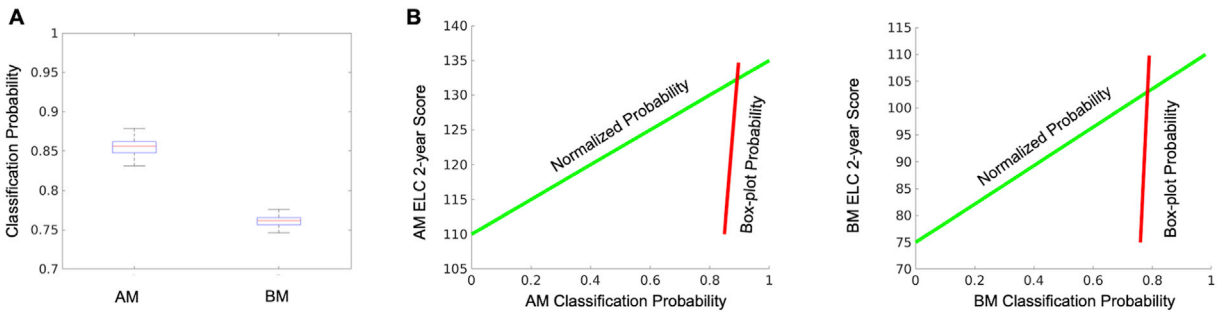

Fig. 3. Two-step prediction pipeline design. The box plots in (A) show the distribution of classification probability values for infants assigned by the model to the correct median score group. Since the variability of the underlying data distribution is small this translates to a very large learning step as seen by the red lines in (B), which in turn may be difficult for a single model design. Alternatively, by introducing a second model building step, which normalizes the classification probabilities, the large learning step is transformed into a smaller learning step as seen by the green lines in (B). As a result, a two-step prediction pipeline design approach may better identify subtle correlations between white matter connectomes and cognitive ability.

the mean absolute error to quantify the difference in score points between actual and predicted ELC scores. Additional information on the prediction model can be found in Supplement Section IV.

2.10. Application to a preterm dataset

To test how well our ELC 2-year prediction pipeline trained with FT infant data generalizes to new infant connectome datasets, we evaluated its accuracy when applied to fully independent datasets from PT infants scanned at approximate term-age-equivalent. This allows us to determine if the trained models are overfit to the FT infant data; achieving similar classification accuracies in a separate dataset of infants with potentially distinct WM characteristics and differential risk for poor cognitive outcomes would suggest that our pipeline models are not overfit and can detect fundamental WM organizational characteristics associated with emergent cognition. Our prediction model is based on the 10-fold cross-validation and thus generates 10 separate predictions for the PT infants, which we average for an overall prediction. More detailed information regarding the PT infant analysis can be found in Supplement Section III.C and Section IV.C.

2.11. Extraction of the connectivity fingerprint

A backtrack approach similar to that proposed in Hazlett et al. (2017b) was employed to identify a systematic WM connectivity pattern, or connectivity fingerprint, important for classification accuracy. In general, our approach works backwards through layers of the trained neural network to the input layer and follows the nodes that have the largest contribution to the layer directly above. When the backtrack approach completes, each WM connection is assigned a normalized weight that indicates its contribution to classification accuracy. Next, using the assigned weight values, the connectivity fingerprint is formed by including only those WM connections that account for the top 20% of the total weight. In the results, we present this backtrack weight per connection in the fingerprint. Brain regions are also assigned a *frequency* that is based on the number of times they appear in separate connections (i.e. if connections between regions A and B and regions A and C are identified through the backtrack approach, region A will have a frequency of 2). More detailed information on the backtrack analysis can be seen in Supplement Section V.

3. Results

3.1. Classification of ELC score group

Classification labels were generated by grouping infants into two categories based on their performance on the 2-year ELC relative to others in our sample: above the median (AM; ELC > 110; FT: $n = 38$, PT: $n = 15$) and below the median (BM; ELC < 110; FT: $n = 37$, PT: $n = 22$). The ELC median score group classification model achieved 89.5% ($SD = 5.7\%$, SE = 0.15%) accuracy for the FT infants and 83.8% accuracy for PT infants.

3.2. Prediction of ELC scores

Plots comparing the predicted ELC score to actual scores can be seen in Fig. 4A for FT infants and Fig. 4B for PT infants. Correlations between predicted and actual scores were high for both FT ($r = 0.978$, $df = 73$, $p < 0.0001$) and PT infants ($r = 0.956$, $df = 35$, $p < 0.0001$; for mean predicted scores across 10 folds). The mean absolute error across individuals for the prediction was 3.45 points (SD = 1.48) for FT and 4.47 points (SD = 1.56) for PT infants.

3.3. Connectivity fingerprint

The connectivity fingerprint shown in Fig. 5 defines region-to-region WM connections that are important for accurate classification based on ELC score group. WM connections contributing to classification accuracy (Fig. 5A) span the brain, with the majority of connections within the frontal cortex and between regions in the frontal lobe and other brain areas including the occipital, temporal, and parietal cortices, insula, cingulate, and pre- and postcentral gyri. Additional connections important for classification include those between the cingulum cortex and the occipital, temporal, and parietal cortices, as well as connections between the occipital cortex and parietal and temporal cortices. Connections with the largest backtrack weight were inter-hemispheric connections in the frontal lobe, including connections between the right superior frontal and left middle frontal gyri, left superior orbitofrontal and right inferior orbitofrontal gyri, and connections between the left olfactory gyrus and the right precentral gyrus, right middle frontal gyrus, and right frontal inferior triangularis. Connections within the right hemisphere and inter-hemispheric connections each account for 40% of backtrack connections, while connections within the left hemisphere account for the remaining 20% of backtrack connections.

Brain regions with the highest frequencies among backtrack connections are found in the frontal cortex and include the bilateral frontal superior medial cortex, right orbitofrontal medial cortex, right frontal inferior operculum, right rectus gyrus, and right inferior parietal and occipital gyri (Fig. 5B). Additional brain regions of lower frequency include the left frontal inferior operculum, left olfactory, left orbitofrontal medial cortex, left rectus gyrus, right anterior and left middle cingulate cortices, right cuneus, right superior occipital cortex, left fusiform gyrus, right angular gyrus, and right inferior temporal gyrus.

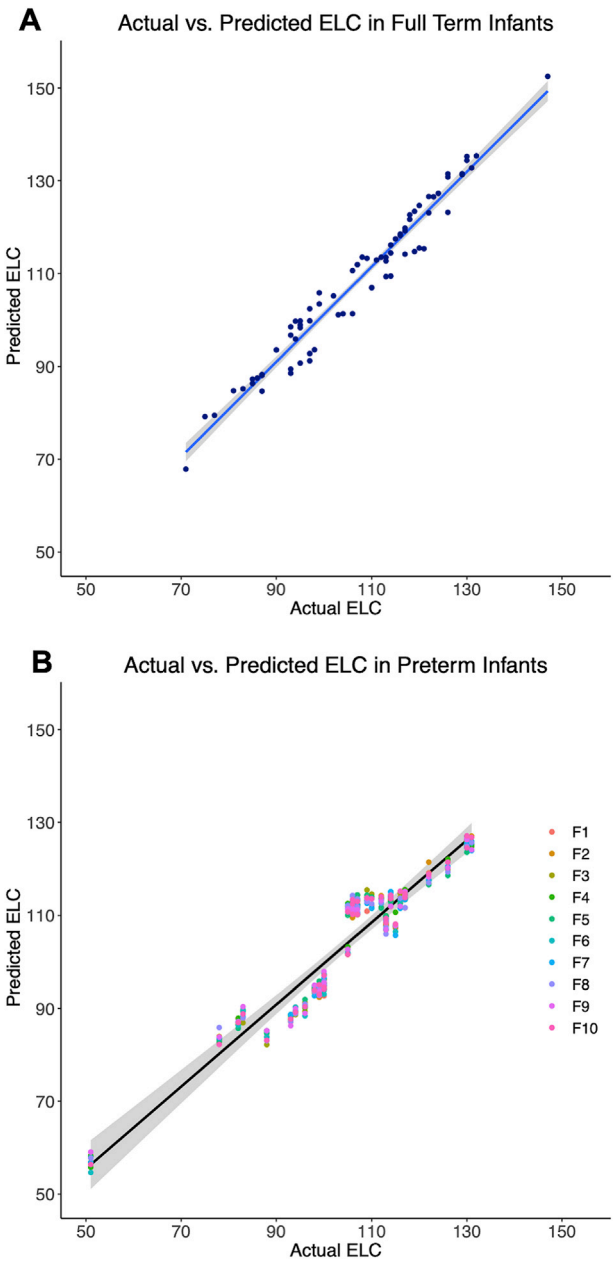

Fig. 4. Correlations between actual and predicted 2-year cognitive scores. Estimated ELC scores generated through regression prediction models are plotted against each FT infant's actual ELC score, along with the linear regression fit line (blue) and shaded 95% confidence intervals (A). PT infants were classified using each of the 10 models generated through cross-validation. Estimated ELC scores are presented for each of the 10 folds (B), where the first fold is represented by F1, along with the linear regression fit line (black) for the mean predicted score across folds and a 95% confidence interval for the regression fit.

4. Discussion

Using a developmental connectomics framework coupled with a deep learning approach, we have demonstrated the ability to accurately predict an infant's cognitive performance at age 2 using WM connectivity matrices generated from scans following birth. Specifically, we show that taking a two-step approach by first classifying infants based on their cognitive performance group and then using results from this group classification to directly predict ELC scores, allowed us to achieve estimates of children's cognitive scores two years later that are highly correlated with their actual scores. Importantly, we found that our prediction model, which was trained using WM data from term-born infants, was equally effective in infants born preterm, despite reported associations between prematurity and altered WM development (Elitt and Rosenberg, 2014). Finally, we report neonatal WM connections that may be particularly important for supporting emergent cognitive abilities at age 2 in both full-term and preterm infants. This study demonstrates the importance of fetal and early postnatal brain development for subsequent cognition and suggests that the WM connectome is an important biomarker of cognitive abilities in early life that deserves further study.

4.1. WM connectomes at birth as biomarkers of cognition in infancy

Neuroimaging research has greatly improved our understanding of human brain function and development; however, it has not yet propelled significant changes in clinical or educational practice (Gabrieli et al., 2015). Thus, as the field advances, there is a critical need to identify neuroimaging biomarkers, or 'neuromarkers' that can aid in improving diagnostic criteria and identifying individuals at risk for poor mental health and cognitive outcomes so that adequate interventions can be designed and implemented. Our study using non-invasive neuromarkers, in this case, WM connectomes, to accurately predict infants' future cognitive performance is an important step in this direction. Machine learning, and particularly deep learning, have been instrumental in neuroimaging research by allowing complex patterns between brain structural and functional features and cognitive and clinical phenotypes to be revealed (Shen et al., 2017; Vieira et al., 2017). In this study, we trained a deep learning classification model to use infant WM connectomes following birth to predict future cognitive outcomes 2 years post-birth. The classification accuracies were relatively high; we achieved 89.5% accuracy in full-term and 83.8% accuracy in preterm infants. This highlights the promise of WM networks as a neuromarker of cognitive abilities during development.

Given the potential clinical relevance of predicting cognitive scores directly, we used classification probabilities from the ELC median score group classification model as inputs in a regression prediction model to estimate each infant's future cognitive score at age 2. This two-step process achieved very high correlations between predicted and actual scores, within just a few points of the actual score in many cases. One preterm infant had a very low score of 51 in the ELC at age 2. Despite scoring 20 points lower than the lowest-scoring full-term infant used to train the model, the infant was accurately classified as below average. While the average predicted score of 57 was above the infant's actual score, this predicted score was still the overall lowest predicted score in both the full-term and preterm samples. Thus, our prediction method correctly identified this subject as the lowest performing subject in the sample, which is an indication of its robustness.

The mean absolute error between predicted and actual scores are, on average, comparable to the standard errors of measurement for the ELC, which range between about 3 and 4 score points for 2-year-olds, and reflect a band of error around the mean, or "true", score that would be obtained if an individual could be tested repeatedly without influences of practice or other factors (Mullen, 1995). This level of accuracy is critical for using neuromarkers to potentially guide cognitive and behavioral interventions in young children.

4.2. Methodological considerations

Through preliminary experimentation we found that a single neural network was not capable of accurately predicting a continuous cognitive measure when applied to a 78×78 dimension connectome (3003 connectivity features) and ~70 FT datasets. This prompted the implementation of a pipeline that defines two sequential models (classification followed by prediction); this was an important step designed to mitigate overfitting by reducing the complexity of the learning task, which is inherent with high dimensional data and limited sample sizes (see

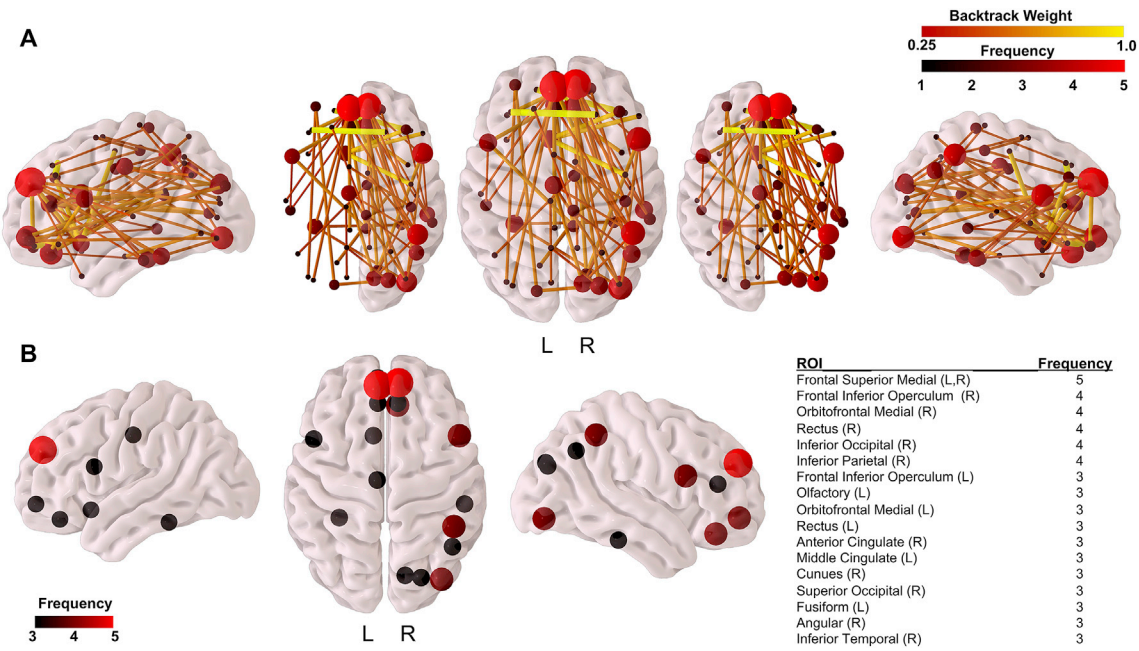

Fig. 5. Connectivity fingerprint of WM connections important for ELC group classification accuracy. Results from the backtrack approach identifying region-to-region connections in the WM connectomes that are most responsible for ELC median score group classification accuracy are shown in (A), where brain regions in this fingerprint are colored by frequency (number of backtrack connections to that brain region) and connections are colored by their backtrack weight [0.25,1]. The highest frequency regions (frequency ≥ 3) are visualized in (B) along with a listing of the anatomical regions from the parcellation atlas.

Supplement Section VI for details).

When compared to previous single-model machine learning approaches that use infant structural connectomes (Kawahara et al., 2017) or adult functional connectomes (Finn et al., 2015) to predict a continuous cognitive measure, our two-step approach is more accurate even though the connectome samples are similar. In particular, the convolutional neural network model developed Kawahara et al. used 90×90 dimension connectomes (4005 connectivity features) and 168 infant connectome samples, and the linear regression model developed by Finn et al. used 268×268 dimension connectomes (35,778 connectivity features) and 118 adult connectome samples. Even though Kawahara et al. applied a procedure to synthetically increase the number of samples, and Finn et al. applied a correlation technique to reduce the number of connectivity features, the highest correlations achieved between predicted and actual scores were $r = 0.31$ and $r = 0.50$, respectively. In general, there could be any number of reasons for the lower correlations (including differences in images and connectome processing), but it is reasonable to assume the high complexity of the machine learning problem resulted in an overfit model that, when applied to unseen data, resulted in lower prediction accuracy than the two-step design which was purposely implemented to minimize such errors in this study.

In addition to the two-step approach, several additional steps have been taken to mitigate overfitting errors in our results: (1) a 10-fold cross-validation strategy was used in which infants in each fold were randomly selected and the outcome ratios were maintained, (2) an integrated 10-fold grid-search capability was performed that identified the optimal model parameter values, and (3) the addition of several dropout layers that further prevent model overfitting (Srivastava et al., 2014) were included in the neural network design. Even though our study has a limited number of datasets, the overfitting precautionary steps we have introduced produce a prediction pipeline that is reliable even when applied to connectomes of preterm infants from an independent data set, which suggests that, if overfitting is indeed happening, the impact is minimal.

4.3. Application to an independent sample of preterm infants

Importantly, our models, which are trained using data from full-term infants, robustly predicted future cognitive scores with similar accuracy using WM connectomes from a sample of preterm infants scanned near term-age-equivalent. The ability to accurately classify across both groups was of clinical interest; preterm birth is associated with alterations in WM microstructural development (Partridge et al., 2005) and increased risk for poor cognitive outcomes (Bode et al., 2014) and neurodevelopmental disorders including attention deficit hyperactivity disorder and autism, which have been linked to disruptions in brain connectivity (Liston et al., 2011; Uddin et al., 2013). Despite the impact of prematurity on neurodevelopment, our model was able to accurately predict cognitive scores in both groups, suggesting there exists an underlying set of organizational principles that govern WM network topology across early pre- and postnatal development and have important implications for cognitive development. This highlights the potential usefulness of WM connectomes as neuromarkers of cognition across heterogeneous infant populations.

4.4. WM connections important for prediction accuracy

A connectivity fingerprint consisting of connections spanning the cortex at birth contributed to classification accuracy, with the highest backtrack-weighted connections confined to the frontal lobe. These findings are consistent with previous reports that rich club regions in the infant WM connectome include regions in the medial frontal cortex (Ball et al., 2014). Interestingly, studies linking infant WM connectomes to parental reports of children's behavior at age 4 also found connectivity to the right inferior frontal gyrus to be an important predictor of future externalizing behavior at 48 months (Wee et al., 2016), and a WM hub in the middle frontal gyrus in preterm infants was found to be related to cognitive abilities at age 2 (Kawahara et al., 2017). These findings, along with the results from our study, suggest that frontal lobe WM connections play an important role in early learning and cognition. We also found that connections between the cingulate and other brain areas were important

for classification accuracy, suggesting that the cingulate, reported to support cognitive control in adults (MacDonald et al., 2000; Shenhav et al., 2016), may also be important for emerging cognition and deserves further study. Finally, many more connections responsible for prediction were located in the right hemisphere than the left, which is in line with work demonstrating a rightward asymmetry of WM network topology in neonates (Yap et al., 2011) as well as adolescents and adults (Zhong et al., 2017). However, the connections with the highest backtrack weight are interhemispheric, and the atlas used in this study is not symmetric, thus we suggest caution in the interpretation of this rightward lateralization.

Results from this study suggest that WM connectomes following birth, as a reflection of fetal brain development, have important implications for future cognitive capacities in children. WM connections in the developing brain are built through genetically regulated cascades of cellular events that govern neurogenesis and migration and promote an exuberance of connections via processes of axon guidance, synaptogenesis and dendritic arborization (Stiles and Jernigan, 2010). In the final trimester, the infant WM connectome undergoes substantial refinement through apoptotic mechanisms that promote the pruning of axons, a process which continues through early postnatal life (Innocenti and Price, 2005). Disruptions in any of these processes may have lasting consequences, as has been suggested by findings that maternal immune activation (Knuesel et al., 2014), maternal stress (Bale, 2015), and maternal drug use (B. L. Thompson et al., 2009) during pregnancy cause alterations in brain development that increase risk for poor cognitive, behavioral, neurological and mental health outcomes in offspring. Our findings would suggest that the foundational wiring of WM circuitry important for future cognition is set in place *in utero* and has a lasting impact on child development, highlighting the importance of understanding how genetic and environmental factors jointly shape fetal brain development.

4.5. Study limitations

This study has limitations that should be considered. First, a cortical-only atlas was used for probabilistic tractography, which does not include cortico-subcortical connections which may be important for emerging cognition, though prediction accuracy was very good without considering subcortical connectivity. Second, cognitive scores in toddlerhood are not particularly strong indicators of later childhood cognitive ability, with 2-year ELC scores accounting for roughly 20% of the variance in 6-year IQ scores, and even weaker associations found for preterm infants (Girault et al., 2018b). Therefore, longitudinal studies will be needed to determine if the neonatal WM connectome serves as a biomarker for later childhood performance. Third, interpreting diffusion connectivity is not trivial; the number of streamlines generated during probabilistic tractography is not a direct measure of anatomical connectivity (Jbabdi and Johansen-Berg, 2011) and partial volume effects present at the spatial resolution of the DWIs used in this study will influence the number of streamlines detected. Finally, findings from this study are limited by small sample sizes and results may not generalize to larger populations, it will be critically important to replicate these findings in other longitudinal datasets, for example, using publicly available data from the Baby Connectome and Developing Human Connectome projects in the US and Europe.

5. Conclusions

Findings from this study indicate that the infant WM connectome is predictive of cognitive performance at age 2, which highlights the importance of WM development *in utero* for subsequent cognition. Our study has implications for screening and intervention and suggests that future work should focus on identifying the ways in which prenatal mechanisms of WM development are influenced by genetic factors as well as the intra- and extrauterine environment to shape individual differences in the WM connectome that serves as a biological foundation for learning.

Acknowledgements and Disclosures

This work was supported by the National Institutes of Health (MH064065, MH070890, MH111944 and HD053000 to JHG, and T32-HD07376 to JBG). We thank our participating families and staff of the UNC Early Brain Development Study, UNC Neuro Image and Research Analysis Laboratory, and Frank Porter Graham Child Development Institute.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.02.060.

References

Avants, B.B., Tustison, N.J., Wu, J., Cook, P.A., Gee, J.C., 2011. An open source multivariate framework for n-tissue segmentation with evaluation on public data. Neuroinformatics 9, 381–400. https://doi.org/10.1007/s12021-011-9109-y.

Bale, T.L., 2015. Epigenetic and transgenerational reprogramming of brain development. Nat. Rev. Neurosci. 16, 332–344. https://doi.org/10.1038/nrn3818.

Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., Robinson, E.C., Ogundipe, E., Rueckert, D., Edwards, A.D., Counsell, S.J., 2014. Rich-club organization of the newborn human brain. Proc. Natl. Acad. Sci. U.S.A. 111, 7456–7461. https://doi.org/10.1073/pnas.1324118111.

Ball, G., Pazderova, L., Chew, A., Tusor, N., Merchant, N., Arichi, T., Allsop, J.M., Cowan, F.M., Edwards, A.D., Counsell, S.J., 2015. Thalamocortical connectivity predicts cognition in children born preterm. Cerebr. Cortex 25, 4310–4318. https://doi.org/10.1093/cercor/bhu331.

Bayley, N., 1949. Consistency and variability in the growth of intelligence from birth to eighteen years. The Pedagogical Seminary and Journal of Genetic Psychology 75, 165–196. https://doi.org/10.1080/08856559.1949.10533516.

Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? Neuroimage 34, 144–155. https://doi.org/10.1016/j.neuroimage.2006.09.018.

Bishop, E.G., Cherny, S.S., Corley, R., Plomin, R., Defries, J.C., Hewitt, J.K., 2003. Development genetic analysis of general cognitive ability from 1 to 12 years in a sample of adoptees, biological siblings, and twins. Intelligence 31, 31–49. https://doi.org/10.1016/S0160-2896(02)00112-5.

Bode, M.M., D'Eugenio, D.B., Mettelman, B.B., Gross, S.J., 2014. Predictive validity of the Bayley, Third Edition at 2 years for intelligence quotient at 4 years in preterm infants. J. Dev. Behav. Pediatr. 35, 570–575. https://doi.org/10.1097/DBP.0000000000000110.

Bradway, K.P., Thompson, C.W., 1962. Intelligence at adulthood: a twenty-five year follow-up. J. Educ. Psychol. 53, 1–14. https://doi.org/10.1037/h0045764.

Cao, M., He, Y., Dai, Z., Liao, X., Jeon, T., Ouyang, M., Chalak, L., Bi, Y., Rollins, N., Dong, Q., Huang, H., 2017a. Early development of functional network segregation revealed by connectomic analysis of the preterm human brain. Cerebr. Cortex 27, 1949–1963. https://doi.org/10.1093/cercor/bhw038.

Cao, M., Huang, H., He, Y., 2017b. Developmental connectomics from infancy through early childhood. Trends Neurosci. 40, 494–506. https://doi.org/10.1016/j.tins.2017.06.003.

Choi, H., Jin, K.H., 2016. Fast and robust segmentation of the striatum using deep convolutional neural networks. J. Neurosci. Methods 274, 146–153. https://doi.org/10.1016/j.jneumeth.2016.10.007.

Cohen, J.R., D'Esposito, M., 2016. The segregation and integration of distinct brain networks and their relationship to cognition. J. Neurosci. 36, 12083–12094. https://doi.org/10.1523/JNEUROSCI.2965-15.2016.

Deary, I.J., Pattie, A., Starr, J.M., 2013. The stability of intelligence from age 11 to age 90 years: the Lothian birth cohort of 1921. Psychol. Sci. 24, 2361–2368. https://doi.org/10.1177/0956797613486487.

Deary, I.J., Whiteman, M.C., Starr, J.M., Whalley, L.J., Fox, H.C., 2004. The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. J. Pers. Soc. Psychol. 86, 130–147. https://doi.org/10.1037/0022-3514.86.1.130.

Dickson, H., Laurens, K.R., Cullen, A.E., Hodgins, S., 2012. Meta-analyses of cognitive and motor function in youth aged 16 years and younger who subsequently develop schizophrenia. Psychol. Med. 42, 743–755. https://doi.org/10.1017/S0033291711001693.

Dubois, J., Dehaene-Lambertz, G., Kulikova, S., Poupon, C., 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.

Dubois, J., Dehaene-Lambertz, G., Perrin, M., Mangin, J.-F., Cointepas, Y., Duchesnay, E., Le Bihan, D., Hertz-Pannier, L., 2008. Asynchrony of the early maturation of white matter bundles in healthy infants: quantitative landmarks revealed noninvasively by diffusion tensor imaging. Hum. Brain Mapp. 29, 14–27. https://doi.org/10.1002/hbm.20363.

Elitt, C.M., Rosenberg, P.A., 2014. The challenge of understanding cerebral white matter injury in the premature infant. Neuroscience 276, 216–238. https://doi.org/10.1016/j.neuroscience.2014.04.038.

Emerson, R.W., Adams, C., Nishino, T., Hazlett, H.C., Wolff, J.J., Zwaigenbaum, L., Constantino, J.N., Shen, M.D., Swanson, M.R., Elison, J.T., Kandala, S., Estes, A.M., Botteron, K.N., Collins, L., Dager, S.R., Evans, A.C., Gerig, G., Gu, H., McKinstry, R.C., Paterson, S., Schultz, R.T., Styner, M., IBIS Network, Schlaggar, B.L., Pruett, J.R., Piven, J., 2017. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. Sci. Transl. Med. 9 https://doi.org/10.1126/scitranslmed.aag2882 eaag2882.

Essen, D.C.V., 1997. A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature 385, 313–318. https://doi.org/10.1038/385313a0.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671. https://doi.org/10.1038/nn.4135.

Fransson, P., Åden, U., Blennow, M., Lagercrantz, H., 2010. The functional architecture of the infant brain as revealed by resting-state fMRI. Cerebr. Cortex 21, 145–154. https://doi.org/10.1093/cercor/bhq071.

Gabrieli, J.D.E., Ghosh, S.S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. Neuron 85, 11–26. https://doi.org/10.1016/j.neuron.2014.10.047.

Gale, C.R., Deary, I.J., Boyle, S.H., Barefoot, J., Mortensen, L.H., Batty, G.D., 2008. Cognitive ability in early adulthood and risk of 5 specific psychiatric disorders in middle age: the Vietnam experience study. Arch. Gen. Psychiatr. 65, 1410–1418. https://doi.org/10.1001/archpsyc.65.12.1410.

Geng, X., Gouttard, S., Sharma, A., Gu, H., Styner, M., Lin, W., Gerig, G., Gilmore, J.H., 2012. Quantitative tract-based white matter development from birth to age 2 years. Neuroimage 61, 542–557. https://doi.org/10.1016/j.neuroimage.2012.03.057.

Gilmore, J.H., Knickmeyer, R.C., Gao, W., 2018. Imaging structural and functional brain development in early childhood. Nat. Rev. Neurosci. 19, 123–137. https://doi.org/10.1038/nrn.2018.1.

Gilmore, J.H., Schmitt, J.E., Knickmeyer, R.C., Smith, J.K., Lin, W., Styner, M., Gerig, G., Neale, M.C., 2010. Genetic and environmental contributions to neonatal brain structure: a twin study. Hum. Brain Mapp. 31, 1174–1182. https://doi.org/10.1002/hbm.20926.

Gilmore, J.H., Shi, F., Woolson, S.L., Knickmeyer, R.C., Short, S.J., Lin, W., Zhu, H., Hamer, R.M., Styner, M., Shen, D., 2012. Longitudinal development of cortical and subcortical gray matter from birth to 2 years. Cerebr. Cortex 22, 2478–2485. https://doi.org/10.1093/cercor/bhr327.

Girault, J.B., Cornea, E., Goldman, B.D., Knickmeyer, R.C., Styner, M., Gilmore, J.H., 2018a. White matter microstructural development and cognitive ability in the first 2 years of life. Hum. Brain Mapp. 111, 7456. https://doi.org/10.1002/hbm.24439.

Girault, J.B., Langworthy, B.W., Goldman, B.D., Stephens, R.L., Cornea, E., Steven Reznick, J., Fine, J., Gilmore, J.H., 2018b. The predictive value of developmental assessments at 1 and 2 for intelligence quotients at 6. Intelligence 68, 58–65. https://doi.org/10.1016/j.intell.2018.03.003.

Gunnell, D., Harrison, G., Rasmussen, F., Fouskakis, D., Tynelius, P., 2002. Associations between premorbid intellectual performance, early-life exposures and early-onset schizophrenia. Cohort study. Br. J. Psychiatry 181, 298–305.

Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V.J., Meuli, R., Thiran, J.P., Grant, P.E., 2010. White matter maturation reshapes structural connectivity in the late developing human brain. Proc. Natl. Acad. Sci. U.S.A. 107, 19067–19072. https://doi.org/10.1073/pnas.1009073107.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N., Collins, D.L., Constantino, J.N., Dager, S.R., Estes, A.M., Evans, A.C., Fonov, V.S., Gerig, G., Kostopoulos, P., McKinstry, R.C., Pandey, J., Paterson, S., Pruett, J.R., Schultz, R.T., Shaw, D.W., Zwaigenbaum, L., Piven, J., IBIS Network, Clinical Sites, Data Coordinating Center, Image Processing Core, Statistical Analysis, 2017a. Early brain development in infants at high risk for autism spectrum disorder. Nature 542, 348–351. https://doi.org/10.1038/nature21369.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N., Collins, D.L., Constantino, J.N., Dager, S.R., Estes, A.M., Evans, A.C., Fonov, V.S., Gerig, G., Kostopoulos, P., McKinstry, R.C., Pandey, J., Paterson, S., Pruett, J.R., Schultz, R.T., Shaw, D.W., Zwaigenbaum, L., Piven, J., IBIS Network, Clinical Sites, Data Coordinating Center, Image Processing Core, Statistical Analysis, 2017b. Early brain development in infants at high risk for autism spectrum disorder. Nature 542, 348–351. https://doi.org/10.1038/nature21369.

Huang, H., Shu, N., Mishra, V., Jeon, T., Chalak, L., Wang, Z.J., Rollins, N., Gong, G., Cheng, H., Peng, Y., Dong, Q., He, Y., 2015. Development of human brain structural networks through infancy and childhood. Cerebr. Cortex 25, 1389–1404. https://doi.org/10.1093/cercor/bht335.

Huang, H., Zhang, J., Wakana, S., Zhang, W., Ren, T., Richards, L.J., Yarowsky, P., Donohue, P., Graham, E., van Zijl, P.C.M., Mori, S., 2006. White and gray matter development in human fetal, newborn and pediatric brains. Neuroimage 33, 27–38. https://doi.org/10.1016/j.neuroimage.2006.06.009.

Innocenti, G.M., Price, D.J., 2005. Exuberance in the development of cortical networks. Nat. Rev. Neurosci. 6, 955–965. https://doi.org/10.1038/nrn1790.

Jbabdi, S., Johansen-Berg, H., 2011. Tractography: where do we go from here? Brain Connect. 1, 169–183. https://doi.org/10.1089/brain.2011.0033.

Jha, S.C., Xia, K., Ahn, M., Girault, J.B., Li, G., Wang, L., Shen, D., Zou, F., Zhu, H., Styner, M., Gilmore, J.H., Knickmeyer, R.C., 2018. Environmental influences on infant cortical thickness and surface area. Cerebr. Cortex 22, 1539. https://doi.org/10.1093/cercor/bhy020.

Kawahara, J., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G., 2017. BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. Neuroimage 146, 1038–1049. https://doi.org/10.1016/j.neuroimage.2016.09.046.

Kim, J., Lee, B., 2018. Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine. Hum. Brain Mapp. 39, 3728–3741. https://doi.org/10.1002/hbm.24207.

Knickmeyer, R.C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J.K., Hamer, R.M., Lin, W., Gerig, G., Gilmore, J.H., 2008. A structural MRI study of human brain development from birth to 2 years. J. Neurosci. 28, 12176–12182. https://doi.org/10.1523/JNEUROSCI.3479-08.2008.

Knickmeyer, R.C., Xia, K., Lu, Z., Ahn, M., Jha, S.C., Zou, F., Zhu, H., Styner, M., Gilmore, J.H., 2016. Impact of demographic and obstetric factors on infant brain volumes: a population neuroscience study. Cerebr. Cortex. https://doi.org/10.1093/cercor/bhw331.

Knuesel, I., Chicha, L., Britschgi, M., Schobel, S.A., Bodmer, M., Hellings, J.A., Toovey, S., Prinssen, E.P., 2014. Maternal immune activation and abnormal brain development across CNS disorders. Nat. Rev. Neurol. 10, 643–660. https://doi.org/10.1038/nrneurol.2014.187.

Koenen, K.C., Moffitt, T.E., Poulton, R., Martin, J., Caspi, A., 2007. Early childhood factors associated with the development of post-traumatic stress disorder: results from a longitudinal birth cohort. Psychol. Med. 37, 181–192. https://doi.org/10.1017/S0033291706009019.

Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R., Heeten, den, A., Karssemeijer, N., 2017. Large scale deep learning for computer aided detection of mammographic lesions. Med. Image Anal. 35, 303–312. https://doi.org/10.1016/j.media.2016.07.007.

Lee, S.J., Steiner, R.J., Yu, Y., Short, S.J., Neale, M.C., Styner, M.A., Zhu, H., Gilmore, J.H., 2017. Common and heritable components of white matter microstructure predict cognitive function at 1 and 2 y. Proc. Natl. Acad. Sci. U.S.A. 114, 148–153. https://doi.org/10.1073/pnas.1604658114.

Li, G., Nie, J., Wang, L., Shi, F., Gilmore, J.H., Lin, W., Shen, D., 2014. Measuring the dynamic longitudinal cortex development in infants by reconstruction of temporally consistent cortical surfaces. Neuroimage 90, 266–279. https://doi.org/10.1016/j.neuroimage.2013.12.038.

Li, G., Nie, J., Wang, L., Shi, F., Lin, W., Gilmore, J.H., Shen, D., 2013. Mapping region-specific longitudinal cortical surface expansion from birth to 2 Years of age. Cerebr. Cortex 23, 2724–2733. https://doi.org/10.1093/cercor/bhs265.

Li, G., Nie, J., Wu, G., Wang, Y., Shen, D., Alzheimer's Disease Neuroimaging Initiative, 2012. Consistent reconstruction of cortical surfaces from longitudinal brain MR images. Neuroimage 59, 3805–3820. https://doi.org/10.1016/j.neuroimage.2011.11.012.

Li, G., Wang, L., Shi, F., Lyall, A.E., Ahn, M., Peng, Z., Zhu, H., Lin, W., Gilmore, J.H., Shen, D., 2016. Cortical thickness and surface area in neonates at high risk for schizophrenia. Brain Struct. Funct. 221, 447–461. https://doi.org/10.1007/s00429-014-0917-3.

Liston, C., Cohen, M.M., Teslovich, T., Levenson, D., Casey, B.J., 2011. Atypical prefrontal connectivity in attention-deficit/hyperactivity disorder: pathway to disease or pathological end point? Biol. Psychiatry 69, 1168–1177. https://doi.org/10.1016/j.biopsych.2011.03.022.

Lyall, A.E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., Hamer, R.M., Shen, D., Gilmore, J.H., 2015. Dynamic development of regional cortical thickness and surface area in early childhood. Cerebr. Cortex 25, 2204–2212. https://doi.org/10.1093/cercor/bhu027.

Mabbott, D.J., Noseworthy, M., Bouffet, E., Laughlin, S., Rockel, C., 2006. White matter growth as a mechanism of cognitive development in children. Neuroimage 33, 936–946. https://doi.org/10.1016/j.neuroimage.2006.07.024.

MacDonald, A.W., Cohen, J.D., Stenger, V.A., Carter, C.S., 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288, 1835–1838.

McCall, R.B., 1977. Childhood IQ's as predictors of adult educational and occupational status. Science 197, 482–483. https://doi.org/10.1126/science.197.4302.482.

McCall, R.B., Hogarty, P.S., Hurlburt, N., 1972. Transitions in infant sensorimotor development and the prediction of childhood IQ. Am. Psychol. 27, 728–748. https://doi.org/10.1037/h0033148.

Mullen, E.M., 1995. Mullen Scales of Early Learning, AGS Edition. American Guidance Service, Inc., Circle Pines, MN.

Nagy, Z., Westerberg, H., Klingberg, T., 2004. Maturation of white matter is associated with the development of cognitive functions during childhood. J. Cognit. Neurosci. 16, 1227–1233. https://doi.org/10.1162/0898929041920441.

Park, H.J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. Science 342. https://doi.org/10.1126/science.1238411, 1238411–1238411.

Partridge, S.C., Mukherjee, P., Berman, J.I., Henry, R.G., Miller, S.P., Lu, Y., Glenn, O.A., Ferriero, D.M., Barkovich, A.J., Vigneron, D.B., 2005. Tractography-based quantitation of diffusion tensor imaging parameters in white matter tracts of preterm newborns. J. Magn. Reson. Imag. 22, 467–474. https://doi.org/10.1002/jmri.20410.

Puechmaille, D., Styner, M., Prieto, J.C., 2017. CIVILITY: cloud based interactive visualization of tractography brain connectome. In: Krol, A., Gimi, B. (Eds.), Presented at the SPIE Medical Imaging. SPIE, p. 101370R. https://doi.org/10.1117/12.2254673.

Shen, D., Wu, G., Suk, H.-I., 2017. Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. 19, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442.

Shenhav, A., Cohen, J.D., Botvinick, M.M., 2016. Dorsal anterior cingulate cortex and the value of control. Nat. Neurosci. 19, 1286–1291. https://doi.org/10.1038/nn.4384.

Sporns, O., 2013. Network attributes for segregation and integration in the human brain. Curr. Opin. Neurobiol. 23, 162–171. https://doi.org/10.1016/j.conb.2012.11.015.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Stiles, J., Jernigan, T.L., 2010. The basics of brain development. Neuropsychol. Rev. 20, 327–348. https://doi.org/10.1007/s11065-010-9148-4.

Stumm, von S., Gale, C.R., Batty, G.D., Deary, I.J., 2009. Childhood intelligence, locus of control and behaviour disturbance as determinants of intergenerational social mobility: British Cohort Study 1970. Intelligence 37, 329–340. https://doi.org/10.1016/j.intell.2009.04.002.

Suk, H.-I., Lee, S.-W., Shen, D., 2015. Alzheimer's Disease Neuroimaging Initiative, 2015. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. Brain Struct. Funct. 220, 841–859. https://doi.org/10.1007/s00429-013-0687-3.

Thompson, B.L., Levitt, P., Stanwood, G.D., 2009. Prenatal exposure to drugs: effects on brain development and implications for policy and education. Nat. Rev. Neurosci. 10, 303–312. https://doi.org/10.1038/nrn2598.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289. https://doi.org/10.1006/nimg.2001.0978.

Uddin, L.Q., Supekar, K., Menon, V., 2013. Reconceptualizing functional brain connectivity in autism from a developmental perspective. Front. Hum. Neurosci. 7 https://doi.org/10.3389/fnhum.2013.00458.

van den Heuvel, M.P., Kersbergen, K.J., de Reus, M.A., Keunen, K., Kahn, R.S., Groenendaal, F., de Vries, L.S., Benders, M.J.N.L., 2015. The neonatal connectome during preterm brain development. Cerebr. Cortex 25, 3000–3013. https://doi.org/10.1093/cercor/bhu095.

Vieira, S., Pinaya, W.H.L., Mechelli, A., 2017. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. Neurosci. Biobehav. Rev. 74, 58–75. https://doi.org/10.1016/j.neubiorev.2017.01.002.

Wee, C.-Y., Tuan, T.A., Broekman, B.F.P., Ong, M.Y., Chong, Y.-S., Kwek, K., Shek, L.P., Saw, S.-M., Gluckman, P.D., Fortier, M.V., Meaney, M.J., Qiu, A., 2016. Neonatal neural networks predict children behavioral profiles later in life. Hum. Brain Mapp. 38, 1362–1373. https://doi.org/10.1002/hbm.23459.

Yap, P.-T., Fan, Y., Chen, Y., Gilmore, J.H., Lin, W., Shen, D., 2011. Development Trends of White Matter Connectivity in the First Years of Life. PLoS One 6, e24678. https://doi.org/10.1371/journal.pone.0024678.

Zammit, S., Allebeck, P., David, A.S., Dalman, C., Hemmingsson, T., Lundberg, I., Lewis, G., 2004. A longitudinal study of premorbid IQ Score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. Arch. Gen. Psychiatr. 61, 354–360. https://doi.org/10.1001/archpsyc.61.4.354.

Zatorre, R.J., Fields, R.D., Johansen-Berg, H., 2012. Plasticity in gray and white: neuroimaging changes in brain structure during learning. Nat. Neurosci. 15, 528–536. https://doi.org/10.1038/nn.3045.

Zhong, S., He, Y., Shu, H., Gong, G., 2017. Developmental changes in topological asymmetry between hemispheric brain white matter networks from adolescence to young adulthood. Cerebr. Cortex 27, 2560–2570. https://doi.org/10.1093/cercor/bhw109.