



## CLINICAL RESEARCH ARTICLE

# Neonatal encephalopathy prediction of poor outcome with diffusion-weighted imaging connectome and fixel-based analysis

Jeong-Won Jeong<sup>1,2,3,4</sup>, Min-Hee Lee<sup>1,2</sup>, Nithi Fernandes<sup>2</sup>, Saihaj Deol<sup>5</sup>, Swati Mody<sup>6</sup>, Suzan Arslanturk<sup>7</sup>, Ratna B. Chinnam<sup>8</sup> and Sidhartha Tan<sup>2</sup>

**BACKGROUND:** Better biomarkers of eventual outcome are needed for neonatal encephalopathy. To identify the most potent neonatal imaging marker associated with 2-year outcomes, we retrospectively performed diffusion-weighted imaging connectome (DWIC) and fixel-based analysis (FBA) on magnetic resonance imaging (MRI) obtained in the first 4 weeks of life in term neonatal encephalopathy newborns.

**METHODS:** Diffusion tractography was available in 15 out of 24 babies with MRI, five each with normal, abnormal motor outcome, or death. All 15 except one underwent hypothermia as initial treatment. In abnormal motor and death groups, DWIC found 19 white matter pathways with severely disrupted fiber orientation distributions.

**RESULTS:** Using random forest classification, these disruptions predicted the follow-up outcomes with 89–99% accuracy. These pathways showed reduced integrity in abnormal motor and death vs. normal tone groups ( $p < 10^{-6}$ ). Using ranked supervised multi-view canonical correlation and depicting just three of the five dimensions of the analysis, the abnormal motor and death were clearly differentiated from each other and the normal tone group.

**CONCLUSIONS:** This study suggests that a machine-learning model for prediction using early DWIC and FBA could be a possible way of developing biomarkers in large MRI datasets having clinical outcomes.

*Pediatric Research* (2022) 91:1505–1515; <https://doi.org/10.1038/s41390-021-01550-2>

**IMPACT:**

- Early connectome and FBA of clinically acquired DWI provide a new noninvasive imaging tool to predict the long-term motor outcomes after birth, based on the severity of white matter injury.
- Disrupted white matter connectivity as a novel neonatal marker achieves high accuracy of 89–99% to predict 2-year motor outcomes using conventional machine-learning classification.
- The proposed neonatal marker may allow better prognostication that is important to elucidate neural repair mechanisms and evaluate treatment modalities in neonatal encephalopathy.

**INTRODUCTION**

In term newborn infants with hypoxic–ischemic encephalopathy (HIE) who underwent cooling, neurodevelopmental performance was predicted by magnetic resonance imaging (MRI) determination of fractional anisotropy (FA) decrease in the white matter on early diffusion tensor imaging (DTI).<sup>1</sup> Basal ganglia and thalamic lesions associated with the severity of motor impairment and abnormal posterior limb of internal capsule signal intensity predicted the inability to walk independently by 2 years.<sup>2</sup> Recent technical improvements in high-field MRI have made abnormalities more detectable. In current clinical practice, MRI interpretations are made by radiologists who may differ in their opinion of what constitutes critical brain injury.

Furthermore, the state of art on prognostication based on MRI findings is still in its early stages. We may need a paradigm shift in achieving prognostication, which is not based on the subjective opinion of a radiologist.

The present study is part of an evaluation of postnatal MRI dependent on “intelligent data mining.” We have recently formed a Big Data group to analyze clinical MRIs in the neonatal period to discover new biomarkers for eventual neurobehavioral outcomes. Indeed, it is of clinical importance to ascertain which aspects of early brain development are predominantly related to the long-term consequences in order to improve early therapeutic interventions for newborn infants with HIE. Our immediate hypothesis was to determine whether the network analysis of

<sup>1</sup>Translational Imaging Laboratory, Children’s Hospital of Michigan, Wayne State University, Detroit, MI, USA; <sup>2</sup>Department of Pediatrics, Wayne State University, Detroit, MI, USA; <sup>3</sup>Department of Neurology, Wayne State University, Detroit, MI, USA; <sup>4</sup>Translational Neuroscience Program, Wayne State University, Detroit, MI, USA; <sup>5</sup>Department of Psychology, Wayne State University, Detroit, MI, USA; <sup>6</sup>Department of Radiology, Wayne State University, Detroit, MI, USA; <sup>7</sup>Department of Computer Science, Wayne State University, Detroit, MI, USA and <sup>8</sup>Department of Industrial and Systems Engineering, Wayne State University, Detroit, MI, USA  
Correspondence: Jeong-Won Jeong (jjeong@med.wayne.edu)

Received: 10 December 2020 Revised: 1 April 2021 Accepted: 8 April 2021  
Published online: 8 May 2021

the whole-brain connectome provides an accurate prediction for 2-year outcomes such as normal tone, abnormal tone, and death.

In this study, we propose the use of clinical diffusion-weighted imaging (DWI) to investigate novel imaging markers at two anatomical levels, (1) axonal pathways using DWI connectome (DWIC)<sup>3,4</sup> and (2) subcortical clusters using fixel-based analysis (FBA).<sup>5</sup> The advantage of such powerful imaging tools is to gauge atypical changes in fiber orientation distribution (FOD) of DWI tractography data reflecting the severity of white matter injury due to severe perinatal hypoxia-ischemia. DWIC and FBA have been used to estimate macroscopic changes in white matter morphology by measuring the count of the streamline tract that connects every pair of cortical regions and the intra-axonal volume measured at individual fiber bundles within every voxel (called "fixel"). This is a feasibility study to use these methods on existing MRI data to identify efficient markers of perinatal white matter injury for accurate prediction of long-term motor outcome.

The objective of the present study is to mine the most effective DWIC-FBA marker of postnatal DWI that accurately predicts three long-term outcomes: death, normal, and abnormal motor (hypertonia) at 2 years follow-up of individual newborn. Our working hypothesis is that interrogation of the cerebral white matter tracts can serve as a biomarker for abnormal neurodevelopment. The streamline tract count can be a questionable marker of neurological connectivity due to the limited spatial resolution of DWI (e.g., ~2 mm) and low sensitivity of the current tractography method to estimate FOD function from clinically acquired DWI tractography data.<sup>6-8</sup> The innovation of this study is that we utilize these limitations to derive a feasible and physiologically meaningful measure in the streamline tract count that quantifies the degree of atypically developed white mature structure in the HIE-affected brain. We then explore in a blinded fashion whether white matter pathways show atypical patterns of the streamline count in subpopulations with different outcomes (in abnormal and death groups compared to normal outcomes), how the atypical patterns in streamline counts are associated with other diffusivity measures, and which subcortical clusters show atypical changes in FBA metrics using a state-of-the-art machine-learning technique.

**METHODS**

**Subjects**

We performed a retrospective study of 24 term newborns of HIE whose neurobehavioral outcome was known on a follow-up visit to the Developmental Assessment Clinic of Children’s Hospital of Michigan. Clinical examination by a neonatologist data was available from electronic medical records. Out of the 24 with MRI, we found that DWI tractography scan was included in 15 newborns (gestation age = 39.2 ± 0.9 weeks and postconceptional age at MRI = 40.7 ± 1.3 weeks, Table 1). These 15 newborns were divided into three groups:

- (1) any death postnatal,
- (2) normal neurological examination at 2 years, especially normal tone, based on neontologist examination at the developmental follow-up clinic, and
- (3) any abnormal tone findings with hypertonia on the neurological exam at 2 years.

The “n” was five per group. The present study was approved by the Wayne State University’s Institutional Review Board, and a waiver of written informed consent was obtained to perform the analysis of existing data in our clinical archive.

**MRI acquisition**

All neonatal MRI scans were performed on a 3 T GE-Signa scanner (General Electric Healthcare Technologies, Milwaukee, WI)

**Table 1.** Perinatal characteristics of the study group.

No.	Outcome	GA	Cooling	Sex	LOS	Age	PMA of MRI	Intensity change and involvement on MRI										NICHD Classification				
								T <sub>1</sub> -w	T <sub>2</sub> -w	Ventricle enlarged	Subdural hematoma	Midline shift	Cystic changes	Frontal lobe	Parietal lobe	Occipital lobe	Cerebellar		Thalamus	Basal ganglia	Brainstem	
1	N	39.3	1	F	18	1	40.3	↑	↓	0	1	0	0	0	1	1	0	0	0	0	0	1A
2	N	38.9	1	F	24	2.4	41.3	↑	0	0	0	0	0	0	0	0	0	0	0	0	0	1A
3	N	38.7	1	M	10	1.6	40.3	0	↑	0	0	0	0	0	0	0	0	0	0	1	0	N/A
4	N	38.0	1	M	14	1.1	39.1	↑	0	0	0	0	0	0	0	0	0	0	0	1	0	2B
5	N	40.0	1	F	9	1.3	41.3	↑	0	0	0	0	0	0	0	0	0	0	1	1	0	2A
6	A	39.5	0	M	11	1.3	41.3	↓	↑	1	0	0	0	0	0	0	0	0	0	1	0	N/A
7	A	39.7	1	M	39	3.7	43.4	↑	0	1	0	1	0	0	0	1	0	0	0	0	0	N/A
8	A	38.0	1	M	10	1.1	39.1	↑	↑	0	0	0	1	0	0	0	0	0	1	1	0	N/A
9	A	38.3	1	F	49	1	39.3	↑	↑	0	0	0	0	0	0	0	0	0	0	1	1	N/A
10	A	39.3	1	M	1	0.1	39.4	↑	↑	0	0	0	0	0	0	1	0	0	0	0	0	N/A
11	D	38.9	1	F	20	2.4	41.3	0	↑	1	0	0	0	0	1	1	1	1	1	1	0	N/A
12	D	41.7	1	F	10	0.9	42.6	0	↑	0	0	0	0	0	0	1	0	0	1	1	0	3
13	D	39.1	1	F	7	0.9	40.0	0	↑	0	0	0	0	0	0	1	0	0	1	1	0	3
14	D	39.0	1	M	16	1.3	40.3	0	↑	0	0	0	0	0	0	1	1	1	1	1	0	3
15	D	40.3	1	M	15	1.3	41.6	0	0	0	0	0	0	0	0	1	1	1	1	1	1	N/A

N normal tone, A abnormal tone, D death, M male, F female, 1 yes, 0 no, GA gestation age, LOS length of stay in the hospital, PMA postmenstrual age, NICHD, National Institute of Child Health and Human Development. LOS is expressed in days, age after birth for MRI in weeks, and PMA of MRI in weeks. MRI readings are provided in a tabular form on the right. ↑: hyper-intensity, ↓: hypo-intensity, N/A classification of MRI injury.

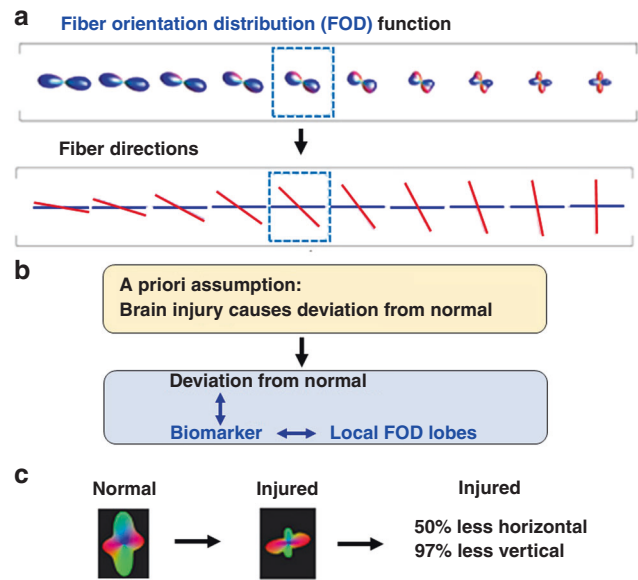
equipped with an 8-channel head coil and ASSET. MRI protocol followed guidelines for routine clinical imaging of DWI, DWI tractography, T<sub>1</sub>-weighted image, and T<sub>2</sub>-weighted image. DWI scan was acquired via echo-planar imaging sequence in the axial plane, with respiratory gating at repetition time/echo time (TR/TE) = 4286/84 ms, field of view (FOV) = 24 cm, 128 × 128 acquisition matrix (nominal resolution = 1.89 mm), contiguous 3 mm thickness to cover entire axial slices of the whole brain. Two *b* values (0 and 700 s/mm<sup>2</sup>) were applied with number of excitations (NEX) = 2. DWI tractography scan was acquired using a double refocusing pulse sequence to reduce eddy current artifacts at TR = 12,500 ms, TE = 88.7 ms, FOV = 24 cm, 128 × 128 acquisition matrix, contiguous 3 mm thickness to cover entire axial slices of whole brain using 33 isotropic gradient directions with *b* = 800 s/mm<sup>2</sup>, 1 *b* = 0 acquisition, and NEX = 1. For morphological analysis, a three-dimensional fast spoiled gradient echo sequence was acquired for the T<sub>1</sub>-weighted sagittal image of each participant at TR/TE/TI of 9.12/3.66/400 ms, slice thickness of 1.2 mm, and planar resolution of 0.78 × 0.78 mm<sup>2</sup>. Axial T<sub>2</sub>-weighted image was acquired with a fast spin-echo sequence, with respiratory gating, at TR/TE of 9231/104 ms (effective) with 5 mm slice thickness, 0 mm gap, FOV = 20 cm, matrix size = 512 × 512, and NEX = 2.

A multidisciplinary team (nurse, neonatal nurse practitioner, and MRI technicians) worked to minimize motion artifacts by a bundle-and-feed protocol, improve the quality of the image acquisition, and allow longer scan time for multiple trials. To minimize the potential confound from motion artifact, the present study excluded patients with unsuccessful MRI showing head motion ≥ 2 mm in DWI encoding data (i.e., voxel size of DWI image), which was evaluated by NIH TORTOISE DWI motion artifact correction package (<https://tortoise.nibib.nih.gov/>).

#### Advantages of extraction and evaluation of DWIC marker

Before performing the tractography analysis, we utilized the NIH TORTOISE DIFF\_PREP package<sup>9</sup> to correct motion, noise, physiological artifacts, susceptibility-induced distortion, and eddy current-induced distortion. FOD function<sup>10,11</sup> was estimated at every voxel of DWI b0 image by using constrained spherical deconvolution (CSD) method<sup>12</sup> that seeks the optimal combinations of multiple fiber compartments in directions and magnitudes of multiple crossing lobe pairs. In contrast to DTI, CSD can model multiple crossing fiber compartments at every single voxel. One hundred dynamically randomized seeding points and angular deviation ≤ 70° were applied at every voxel of the whole brain to reconstruct continuous fiber tract streamlines using the MRtrix3 package (<http://www.mrtrix.org/>) where the second-order integration over FOD (iFOD2) tractography<sup>13</sup> was applied to reconstruct the fiber tract streamlines continuously connecting the most probable neighborhood peak of FOD lobes at every voxel.<sup>14</sup> The advantages of using FOD lobes (Fig. 1a) were principally threefold. Immediately, one could obtain an idea of the directionality and magnitude of crossing fibers involved in a particular fiber tract (Fig. 1a). Second, we made an a priori assumption that brain injury will cause the FOD lobes to deviate from the values in normal regions found in the framework of the conventional tractography method, and result in more spurious streamlines. These deviations can then be used as a potential biomarker (Fig. 1b). Third, when the injury occurs to the fiber tract, the magnitude of the insult can be gauged by estimating abnormal FOD functions and reconstructing the directionality and magnitude of crossing fibers at the HIE-affected region (Fig. 1c).

Whole-brain tracts of individual newborns were characterized by using automated anatomical labeling (AAL) parcellation of UNC neonate atlas<sup>15</sup> that consists of a set of 90 nodes,  $\Omega_i = j = 1-90$  (Fig. 2 and Supplementary Table 1), resulting in whole-brain connectome,  $G = (\Omega, S)$ , where the elements of edges  $S(i, j)$  quantify the pair-wise connectivity strengths between  $\Omega_i$  and  $\Omega_j$  (i.e., the number of fiber streamlines scaled by the total volume of two



**Fig. 1** Fiber orientation distribution (FOD) function as a potential biomarker of white matter injury. **a** Advantage of using fiber orientation distribution (FOD) function is that the crossing fiber compartments simulated in fiber directions can be depicted in the FOD lobe to reflect the contribution of the crossing fibers to the orientation distribution function. **b** A priori assumption of the present study using the local FOD lobes as a potential biomarker. **c** The magnitude of the injury can be depicted in the magnitudes of an example comparing a normal and injured FOD function lobe.

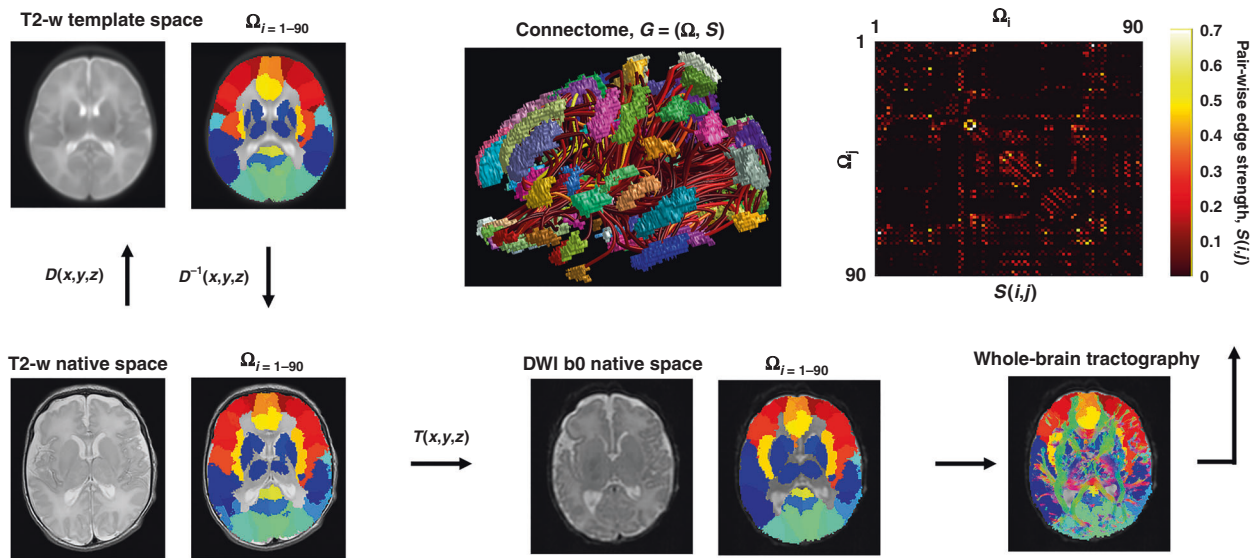
nodes to stabilize intersubject variability by correcting for intracranial volume). Three separate Wilcoxon rank-sum tests: (1) normal vs. abnormal motor, (2) normal motor vs. death, and (3) abnormal motor vs. death at  $p < 0.05$  after Šidák correction for multiple comparisons<sup>16</sup> were then combined to select pair-wise connection edges,  $S(i, j)$  of which log-strengths are significantly altered in three different groups, yielding a marker of  $\log(S(i, j))$  that can quantify significant changes in multiple edge strengths in the whole brain.

In addition to the marker of  $\log(S(i, j))$ , four different markers were created by averaging four diffusivity measures:<sup>17</sup> apparent diffusion coefficient (ADC, the degree of isotropic water diffusion), FA (the degree of white matter integrity), axial diffusivity (AD), and radial diffusivity (RD) at individual streamline tracts included in each edge of  $\log(S(i, j))$  (i.e., streamline tracts connecting  $\Omega_i$  and  $\Omega_j$ ). Combinations of axonal loss and myelin changes may affect combinations of AD and RD, although these may not be precise reflections.<sup>18</sup> We define a dimension as the total range of values obtained in each DWIC edge that has been deemed significant by a priori statistical criteria mentioned above.

#### Extraction and evaluation of FBA marker

FBA has been used to obtain more comprehensive markers reflecting the total number of white matter axons within a voxel.<sup>19-21</sup> Most white matter voxels contain contributions from multiple fiber populations (often referred to as crossing fibers). Therefore, voxel averaged quantitative markers (e.g., FA, AD, RD, ADC, etc.) are not fiber-specific and have poor interpretability. Due to this limitation, fiber density (FD)<sup>5</sup> was estimated as a measure of intra-axonal volume at individual bundles of crossing fibers within every voxel (called “fixels”) by constructing the representative fiber-bundle elements at the group level.

The detailed architecture of FBA has been presented elsewhere.<sup>5,22</sup> This study utilized a pipeline of FBA implemented in the MRtrix3 package (<http://www.mrtrix.org/>). Briefly, a group-average response function was estimated after performing global intensity



**Fig. 2 Schematic of the DWIC analysis to construct the connectome graph,  $G = (\Omega, S)$ , of an individual infant.** Advanced normalization tools (ANTs, <https://github.com/ANTsX/ANTs>) were used to find a 3-D deformation field,  $D(x,y,z)$  that warps T2-w native image into the UNC neonatal T2-w template image (<https://www.med.unc.edu/bric/ideagroup/free-software/unc-infant-0-1-2-atlases/>). The inverse of  $D(x,y,z)$ ,  $D^{-1}(x,y,z)$  was then used to place the UNC neonatal AAL parcellation atlas of 90 cortical nodes,  $\Omega_i = 1-90$ , from T2-w template brain space to T2-w native brain space. Finally, the resulting AAL atlas was placed to native b0 space via nonlinear warping of the ANTs,  $T(x,y,z)$ , between T2-w native image and DWI b0 image and used to sort out whole-brain tracts, leading to an adjacent matrix  $S(i,j)$  of which elements consist of connectivity edge strengths (i.e., the number of fiber streamlines scaled by the total volume of two nodes to stabilize intersubject variability by correcting for intracranial volume). In an example of 3-D visualization (ID #1), colored patches and streamline tubes indicate  $\Omega_i$  and  $S(i,j)$  in the given graph,  $G$ .

normalization across patients and used to reconstruct the FOD functions from 33-direction diffusion data of individual subjects. All FOD images were registered towards a study-specific group-average FOD template ( $n = 15$  patients). Each FOD in the template was segmented into individual fixels by applying a fixel mask at the peak threshold of 0.1, thus defining the position and orientation of all fixels of interest across patients. Warps estimated from registration were applied to deform FOD images to the template space. Warping was done to ensure orientation information remained anatomically consistent across voxels.<sup>22</sup> Each FOD in the warped images was segmented to determine a measure of FD (i.e., FOD lobe integral) per fixel. The estimated FD was compared between groups using fixel-based statistics called threshold-free cluster enhancement (TFCE).<sup>23</sup>

To identify the fixels of which FD values are the most effective in differentiating normal from abnormal motor and death groups, we performed two different TCFE analyses, (1) normal vs. abnormal motor and (2) normal vs. death. The fixel clusters of which FD values differ in two comparisons were obtained at the corrected  $p$  of TFCE  $< 0.05$ . FD values of each cluster were averaged, providing a marker that can quantify overall intra-axonal volume changes at the cluster level of the whole brain. We extracted other two FBA measures: fiber-bundle cross-section (FC), as an estimate of the difference in FC due to the nonlinear warping that transforms FOD functions (or fixels) from subject to template space, and a combined measure of FD and cross-section (FDC), as the multiplication of FD by FC from the same clusters. These are used as additional markers quantifying overall changes in fiber-bundle cross-sectional area and FD-weighted by the difference in the cross-sectional extent of the tract, respectively.

Classification of DWIC and FBA markers for prediction of long-term outcome

It should be noted that this study was originally aimed to investigate whether novel imaging markers underlying the atypically developed

brain abnormalities at two anatomical levels, (1) axonal pathways using DWIC and (2) subcortical clusters using FBA, can provide an accurate prediction for 2-year outcomes. Thus, a set of multiple pathways (or clusters) consists of a multidimensional marker in the feature space. For each marker of DWIC and FBA, an in-house built random forest classification with 100 bagged ensemble of regression trees was used to evaluate individual marker performance in the framework of supervised multiview canonical correlation (SMVCCA).<sup>24,25</sup> The SMVCCA is an iterative process to reduce data dimensionality by fusing or integrating the high multidimensional data into a more amenable data representation for disease classification. It iteratively projects the original data into a given number of eigenvectors of their covariance matrix. In other words, the SMVCCA fuses (or integrates) the multidimensional marker values into lower-dimensional representation to improve separation between clinical outcomes. We defined a “fused dimension” as the given number of eigenvectors and “fused marker” as the projection of the original multidimensional marker values on the given number of eigenvectors, respectively. The steps of SMVCCA are given below:

**Step 1** Iterative data reduction and classification using SMVCCA and Random forest algorithm  $X$ : data matrix,  $X \in R^{N \times M}$ ,  $i \in \{1, \dots, N\}$ ,  $j \in \{1, \dots, M\}$ ,  $N$ : number of subjects,  $M$ : number of multidimensional features;  $Y$ : class label vector,  $Y \in R^{N \times 1}$ ,  $Y_i = 1$  for normal motor, 2 for abnormal motor, 3 for death class

- i. We calculated an optimal weight matrix  $\hat{W} = [W_x \ W_y]$  to maximize feature-feature correlation and class label-feature correlation

$$\operatorname{argmax}_{W_x, W_y} \operatorname{trace}(\hat{W}^T \hat{C} \hat{W}) \text{ s.t. } \hat{W}^T \hat{C}_d \hat{W} = I$$

- ii. The solution of i) is  $\hat{C}_d^{-1} \hat{C} \hat{W} = \hat{W} \Lambda$  that can be solved by  $[\Lambda, D] = \operatorname{eig}(\hat{C}_d, \hat{C})$ , eig: singular value decomposition.

where  $\Lambda$  and  $D$  are the eigenvector matrix and eigen value matrix of  $\hat{C}_d$  and  $\hat{C}$ , respectively.

$$\begin{aligned} \hat{C} &= [\bar{C} X^T Y; Y^T X \text{ zeros (size } (Y^T X, 1), \text{ size } (X^T Y, 2))]; \\ \hat{C}_d &= [\bar{C}_d \text{ zeros (size } (\bar{C}_d, 1), \text{ size } (Y^T Y, 2)); \text{ zeros (size } (Y^T Y, 1), \\ &\text{size } (\bar{C}_d, 2)) Y^T Y]; \\ \bar{C}_d &= \text{zeros (size } (C)); \bar{C}_d(1:(\text{size } (C, 1)+1):\text{end}) = C_d; \bar{C} = C \\ &- \bar{C}_d; \end{aligned}$$

$C$  = covariance matrix of  $X$

- iii. Iterative SMVCCA-based data reduction to fuse high dimensional feature vectors into a low dimensional feature vector,  $f$  and evaluate its classification accuracy in the random forest algorithm:

for  $j = 1: M$

- a. We next created a fused vector,  $f(j)$  with the dimension of  $j$  by using a subset of eigen vectors with the largest eigen values in  $\Lambda(j) = [\Lambda_1, \dots, \Lambda_j]$ ,  $\Lambda_j = j$ th column vector of  $\Lambda$ ,  $W(j) = \Lambda(j)$

$$f(j) = X * \Lambda(j)$$

- b. Then performed the supervised random forest algorithm to classify the fused feature vector,  $f(j)$  into three target classes, normal motor, abnormal motor, and death.
- c. Evaluated the classification accuracy of the fused feature vector,  $f(j)$

$$\text{accuracy}(j) = (\text{True positive} + \text{True negative}) / (\text{True positive} + \text{False positive} + \text{True negative} + \text{False negative})$$

end

**Step 2** The optimal dimension of the fused feature vector,  $\hat{j}$ , maximizing the classification accuracy is determined,

$$\hat{j} = \max_j \text{accuracy}(j)$$

**Step 3** The optimal fused feature vector,  $\hat{f}$ , for outcome prediction then becomes,

$$\hat{f} = X * \Lambda(\hat{j})$$

Using two-fold cross-validation, training data instances ( $X$  and  $Y$  of train set samples) were first used to identify  $\hat{j}$  via Steps 1–3. The identified  $\hat{j}$  was applied to test data instances ( $X$  and  $Y$  of train set samples) in order to assess the performance metrics of SMVCCA for outcome prediction (i.e., accuracy, sensitivity, specificity). This study repeated the above cross-validation 100 times. The mean and standard deviation of the performance metrics over these repetitions were reported in Table 3.

In other words, two-fold cross-validation was applied to split the fused markers of the entire study cohort into training and test sets. For each split, the bagged ensemble of the regression tree (forest) was optimized to yield maximal accuracy of correct classification at the training set (the first fold,  $n = 8$ ). The optimized forest was then applied to predict the class memberships of the test set (the second fold,  $n = 7$ ). One hundred random splits of the 15 samples into training and test sets were repeated to evaluate the overall accuracy of correct classification for the fused marker. As for the explorative comparison, each element value of the original multi-dimensional marker was ranked according to its magnitude (e.g., 1–15 from the highest to the lowest). The resulting ranked multi-dimensional marker was then fused by the SMVCCA

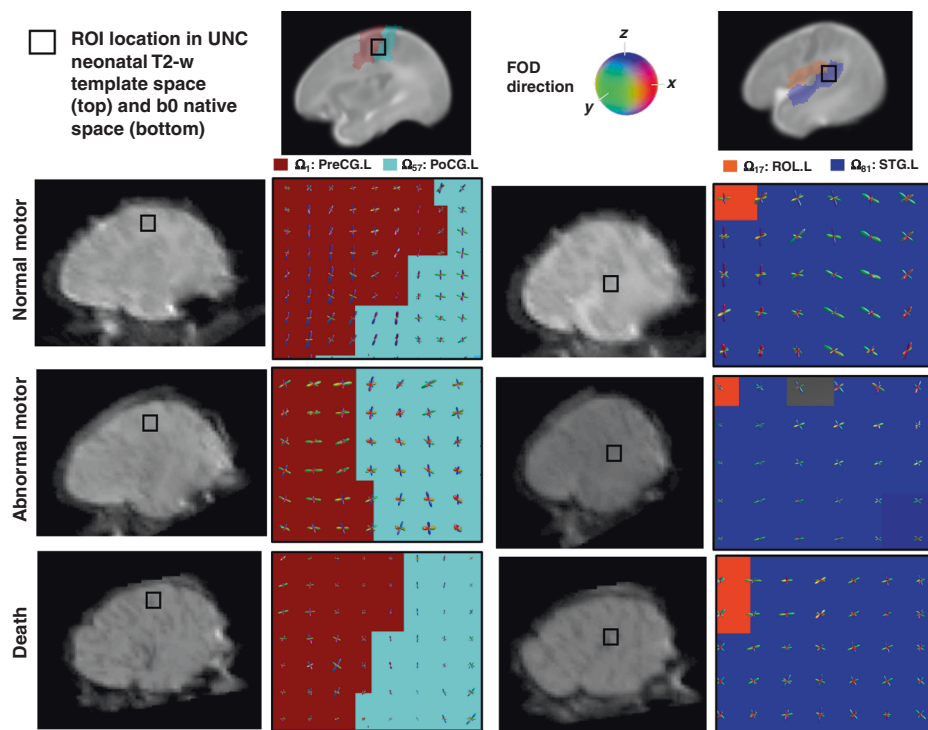
process, and finally re-classified with the forest algorithm using the two-fold cross-validation.

## RESULTS

To demonstrate the feasibility of the iFOD2 method using the second integral of FOD function in our dataset, we estimated the FOD functions at two regions of interest, i.e., the lateral part of the precentral gyrus and superior temporal gyrus, which are the core regions of the primary somatosensory motor system (Fig. 3). In both ROIs, shape and morphological features of FOD functions including magnitude (FOD lobe size), orientation (lobe direction), and the total number of lobes were atypically altered in abnormal motor and death subjects. Lower magnitudes, heterogeneous orientations, and more spurious peaks were found in the two groups compared with normal subjects, implicating injured myelination and disrupted maturation of perinatal white matter in abnormal and death subjects, respectively. In the framework of subsequent probabilistic tractography, these atypical alterations inevitably increased spurious fiber streamlines (i.e., false-positive tracts that do not anatomically exist), leading to paradoxically increased strength of DWIC edge,  $S(i,j)$  in both the abnormal motor and death groups.

Figure 4 presents the results of the proposed DWIC markers determined by Wilcoxon rank-sum tests between three long-term outcome groups. Total 19 pair-wise edges,  $S(i,j)$  (Fig. 4a and Table 2) showed significant difference of group median in their log-strengths,  $\log(S(i,j))$  at the corrected  $p < 0.05$ , yielding a 19-dimensional DWIC marker of  $\log(S(i,j))$ . Because of the overlap in observation in groups and multiple comparisons, caution has to be exercised about the biological importance of the statistical significance. Another way to look at the biological significance is to examine how much the Z-score is beyond the value of  $\pm 1.96$ . Compared with the normal group, pair-wise edge strength,  $\log(S(i,j))$ , was significantly reduced in each of 19 pair-wise pathways in both abnormal tone and death groups with average Z-statistic value =  $-2.552/-2.507$ ,  $p = 0.018/0.012$  for abnormal tone and death, respectively. The overall one-way ANOVAs for  $\log(S(i,j))$  were highly significant even after Sidák correction (a priori  $\alpha = 0.0034$ ) for multiple comparisons ( $p < 10^{-6}$  for normal vs. abnormal, normal vs. death, and abnormal vs. death), indicating the presence of significant differences between groups in the overall dataset. In the comparison of abnormal tone and death groups, we found that compared with the abnormal group, the death group has significantly lower strength in each of these 19 edges, with an average Z-statistic value =  $-2.298$ ,  $p = 0.026$ , which may not be as striking. The group variations of five 19-dimensional DWIC markers,  $\log(S(i,j))$ , FA, AD, RD, and ADC, are shown in Fig. 4b, where each 19-dimensional marker of the individual patient was concatenated per group for two-group comparisons in the box-and-whisker plots. FA showed significant differences in normal vs. abnormal ( $p < 10^{-6}$ ), normal vs. death ( $p < 10^{-6}$ ), and abnormal vs. death ( $p < 10^{-6}$ ). AD showed less striking but still significant differences in normal vs. death ( $p = 0.002$ ). No significant difference was found in AD between in normal vs. abnormal ( $p = 0.016$ ), and abnormal and death groups ( $p = 0.320$ ). RD showed significant differences in normal vs. abnormal ( $p < 10^{-6}$ ), normal vs. death ( $p < 10^{-6}$ ), but not in abnormal vs. death ( $p = 0.026$ ). ADC showed significant differences in normal vs. abnormal ( $p < 0.001$ ) and normal vs. death ( $p < 10^{-6}$ ). No significant difference was found in ADC between abnormal and death groups ( $p = 0.152$ ).

TFCE evaluations of FD (Fig. 5a) found three fixel clusters of interest in thalamus, posterior limb of internal capsule, and cerebellar peduncle, yielding a three-dimensional FD marker for an individual subject. In all three clusters, FOD functions of abnormal and death groups had lower amplitudes leading to lower FD, compared with the normal group. Each element of the



**Fig. 3 To the naked eye, three groups of 2-year outcomes can be differentiated by comparing the shapes of FOD functions in the somatosensory motor system.** Representative examples of fiber orientation distribution (FOD) functions estimated from two regions of interest (ROI) by referring to the UNC neonatal T2-w template images, left lateral portion of the central sulcus and left superior temporal gyrus of three postmenstrual MRI age-matched subjects, normal, abnormal motor and death (all at 1.3 weeks). Left column: FOD functions located in the lateral portion of left central sulcus consisting of two AAL nodes,  $\Omega_1$ : left precentral gyrus (PreCG.L) and  $\Omega_{57}$ : left postcentral gyrus (PoCG.L). Right column: FOD functions located in the lateral portion of left superior temporal gyrus consisting of two AAL nodes,  $\Omega_{17}$ : left Rolandic operculum (ROL.L) and  $\Omega_{81}$ : left superior temporal gyrus (STG.L). Collectively, the total number of lobes, lobe orientation, and lobe size of each FOD function appear to be higher, more inconsistent, and smaller in abnormal motor and death groups, compared with the normal group. A cautionary note is that this phenomenon inevitably increases more spurious fiber streamlines in the framework of probabilistic tractography. Random seeding per voxel continuously reproduces false-positive tracts by tracking false streamlines within a fixed constraint of angular deviation (e.g.,  $\leq 70^\circ$ ).

three-dimensional FD marker (left boxplot of Fig. 5b) suggests a group difference between normal and abnormal ( $p < 0.013$ ). Similarly, each concatenation of the three-dimensional FDC marker samples (right boxplot of Fig. 5b) suggests a group difference between normal and abnormal ( $p < 0.016$ ), with no other differences with  $\alpha < 0.05$  found. These are suggestive findings, because if we use a correction for multiple comparisons, the  $\alpha$  becomes nonsignificant.

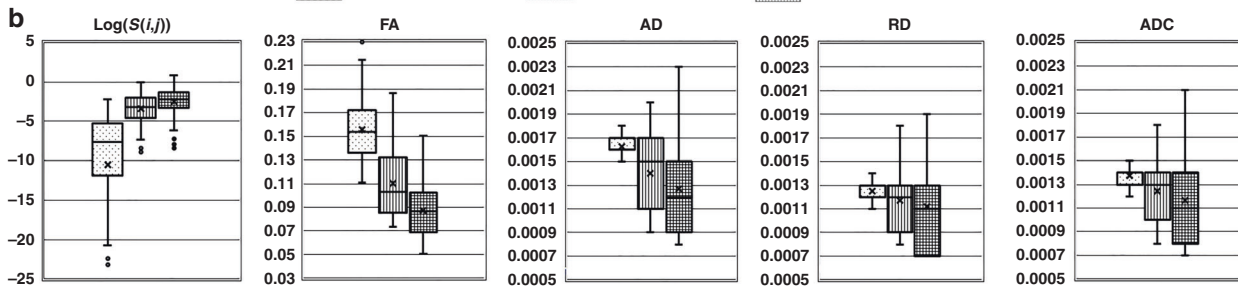
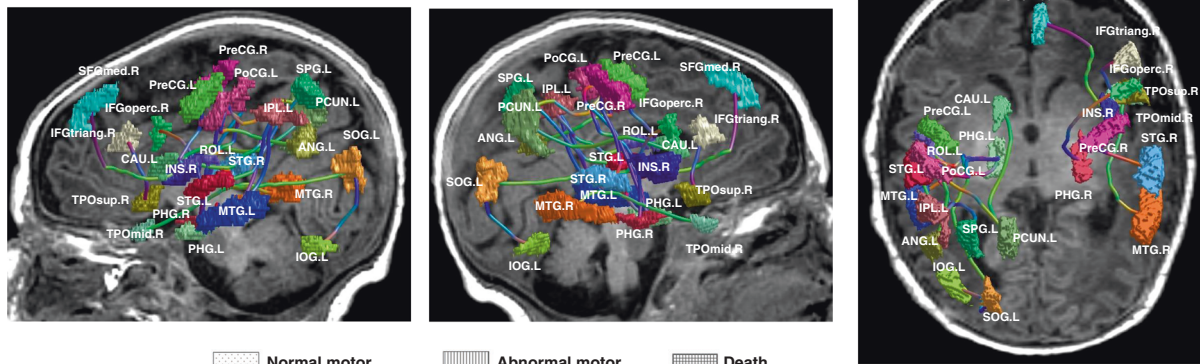
The subsequent random forest classification revealed that compared with other markers including clinical and radiological variables such as sex, gender, gestation age, length of stay in the hospital, intensity change, and involvement on MRI, a DWIC marker of  $\log(S(i,j))$  could achieve the highest accuracy to correctly classify the follow-up motor outcomes, up to 89% without SMVCCA, 92% with SMVCCA, and 99% with ranked SMVCCA (Table 3). Of note, other markers had relatively lower accuracy compared with  $\log(S(i,j))$ , indicating the outperformance of DWIC tract counts to differentiate malformed FOD functions affected by perinatal white matter injuries and immaturities. The  $\log(S(i,j))$  provided the most substantial separation between abnormal tone and death groups. The major finding of this study shows that we fused the 19-dimensional data and got five eigen vectors as the final “fused dimensions” using ranked SMVCCA. We could depict just three of the five “fused dimensions” of the analysis (it is impossible to depict four or five dimensions), and one can visually identify high-risk populations from an MRI study of the eventual outcome (Fig. 6). As denoted by each colored ellipse representing the upper limit of the three fused-dimensional features to predict each group at the confidence level of 99% (i.e.,  $Z\text{-score} = 2.58$

under the assumption of normal distribution), no spatial overlap of the upper limit was found between every pair of three groups at the confidence level of 99%, indicating complete separation of the individual feature to predict three groups in the proposed feature space. This suggests that the combined three fused-dimensional statistic becomes a more powerful biomarker showing high significance than individual DWIC outcomes, which may not reach statistical significance due to multiple comparisons.

## DISCUSSION

The major finding of this study is that a method using SMVCCA can provide a predictive MRI of the possible eventual outcome, in our case, the ranked SMVCCA. The other findings of the present study are paradoxically increased edge strengths ( $\log(S(i,j))$ ), reduced white matter integrity (FA), increased RD, and reduced axonal density (FD) of the multiple white matter pathways in abnormal tone and death groups compared to the control group with normal tone. In the present study, DWIC and FBA of the same dataset were compared to extract the most potent marker for accurate prediction of long-term motor outcomes: normal tone, abnormal tone, and death. We showed perinatal white matter injuries and immaturities in forms of connectivity strength and FD that were altered in the thalamocortical network, including thalamus, posterior limb of the internal capsule, and cerebellar peduncle. This work is the first to look at white matter abnormalities at different formulations of DWI features (i.e., pairwise connection and fiber-specific bundle) that can further improve long-term prediction in the framework of conventional

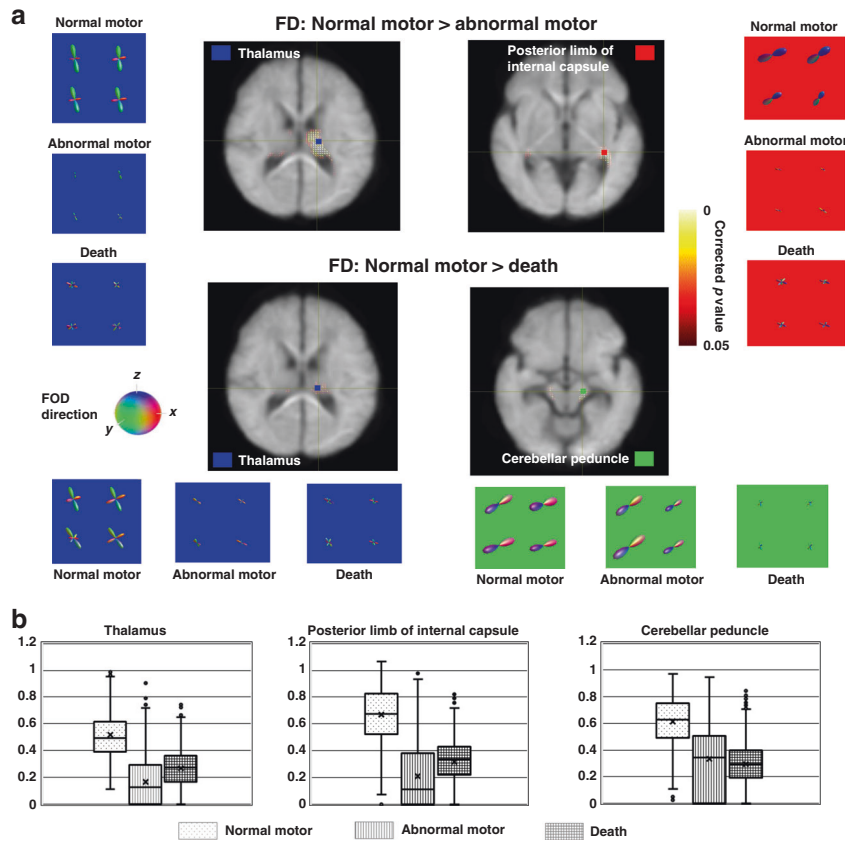
a



**Fig. 4** Trend for differentiating three groups of 2-year outcomes by comparing five DWIC markers that were measured from neural pathways of interest. **a** Total 19 pair-wise connection edges,  $S(i,j)$ , satisfying significant group difference by Wilcoxon rank-sum tests for three comparisons: (1) normal vs. abnormal motor, (2) normal vs. death, and (3) abnormal motor vs. death ( $\alpha < 0.05$ ). Colored patch and streamline tube indicate AAL node,  $\Omega_i$ , and exemplar pathway of  $S(i,j)$  connecting two patches,  $\Omega_i$  and  $\Omega_j$ . Anatomical labels of 90 AAL nodes and two nodes of 19 edges are available in Supplementary Tables 1 and 2. **b** Each diffusivity measure was averaged in all tracts of 19 pair-wise connection edges to define a 19-dimensional marker. Box-and-whisker plots, showing five DWIC markers:  $\log(S(i,j))$ , significantly different for all three comparisons (Šidák correction  $\alpha = 0.0034$  for multiple comparisons); FA, significantly different in all three comparisons; AD, only significant in comparison two (normal vs. death); RD and ADC, in first two comparisons (normal vs. abnormal and vs. death).

**Table 2.** Nineteen pair-wise connection edges,  $S(i,j)$  used to evaluate DWIC markers.

Edge	$\Omega_i$	Label	Anatomy	$\Omega_j$	Label	Anatomy
S(1,17)	1	PreCG.L	Precentral gyrus left	17	ROL.L	Rolandic operculum left
S(2,82)	2	PreCG.R	Precentral gyrus right	82	STG.R	Superior temporal gyrus right
S(10,42)	10	ORBmid.R	Orbitofrontal cortex (middle) right	42	AMYG.R	Amygdala right
S(12,40)	12	IFGoperc.R	Inferior frontal gyrus (opercular) right	40	PHG.R	ParaHippocampal gyrus right
S(14,84)	14	IFGtriang.R	Inferior frontal gyrus (triangular) right	84	TPOsup.R	Temporal pole (superior) right
S(17,39)	17	ROL.L	Rolandic operculum left	39	PHG.L	ParaHippocampal gyrus left
S(17,61)	17	ROL.L	Rolandic operculum left	61	IPL.L	Inferior parietal lobule left
S(24,30)	24	SFGmed.R	Superior frontal gyrus (medial) right	30	INS.R	Insula right
S(39,57)	39	PHG.L	ParaHippocampal gyrus left	57	PoCG.L	Postcentral gyrus left
S(39,85)	39	PHG.L	ParaHippocampal gyrus left	85	MTG.L	Middle temporal gyrus left
S(49,53)	49	SOG.L	Superior occipital gyrus left	53	IOG.L	Inferior occipital gyrus left
S(49,57)	49	SOG.L	Superior occipital gyrus left	57	PoCG.L	Postcentral gyrus left
S(57,65)	57	PoCG.L	Postcentral gyrus left	65	ANG.L	Angular gyrus left
S(57,67)	57	PoCG.L	Postcentral gyrus left	67	PCUN.L	Precuneus left
S(59,85)	59	SPG.L	Superior parietal gyrus left	85	MTG.L	Middle temporal gyrus left
S(61,81)	61	IPL.L	Inferior parietal lobule left	81	STG.L	Superior temporal gyrus left
S(61,85)	61	IPL.L	Inferior parietal lobule left	85	MTG.L	Middle temporal gyrus left
S(67,71)	67	PCUN.L	Precuneus left	71	CAU.L	Caudate left
S(86,88)	86	MTG.R	Middle temporal gyrus right	88	TPOmid.R	Temporal pole (middle) right



**Fig. 5** Using  $\alpha < 0.05$ , significant alterations of FOD functions underlying early hypoxic injuries were found in subcortical regions of abnormal motor and death groups, including thalamus, posterior limb of internal capsule, and cerebellar peduncle. **a** Two TCFE analyses of FD maps, (1) normal vs. abnormal motor and (2) normal vs. death, were performed to identify fixel clusters of interest showing significant deviations from the normal tone group at corrected  $p < 0.05$ . Four clusters were found in two regions of the right thalamus (blue square box), posterior limb of the internal capsule (red square box), and cerebellar peduncle (green square box). In each cluster, abnormal and death groups showed more FOD changes with lobes being narrower and weaker compared with those of the normal tone group, leading to smaller FD values in abnormal motor and death groups. **b** Box-and-whisker plots of FD marker obtained from the voxels of three clusters. Each box indicates the sample range of 25th and 75th percentiles of each group. These are suggestive findings, because if we use a correction for multiple comparisons, the  $\alpha$  becomes nonsignificant.

machine-learning classification. The altered connectivity strength underlying the presence of FOD functions with noisy fiber tract peaks having low lobe amplitudes was confirmed by a significant reduction in white matter integrity (FA), myelination (RD), and axonal volume (FD) measured in the same dataset. The FOD functions with low lobe amplitudes may also reflect inadequate myelination or a state resulting from inadequate crosstalk between axons and oligodendroglia. Our neonatal marker  $\log(S(i,j))$  was gradually increased in abnormal and death groups, yielding a promising accuracy of the correct classification in the framework of conventional machine-learning technique (e.g., 89–99% for the test set, which was not included to train the bagged ensemble of the regression tree). None of the other markers or their combinations, including clinical and radiological variables, differentiated between normal and abnormal tone and death. Please note that the superior performance of  $\log(S(i,j))$  might be explained by the fact that other diffusivity markers were evaluated from streamline tracts of  $S(i,j)$ . A strategy to use other diffusivity measures to preselect streamline tracts of interest (in this study we used streamline count) may not necessarily be better in differentiating between clinical outcomes. Only when combining the various elements of the individual marker using ranked SMVCCA, we found what could be considered a useful biomarker. Our reasoning is that even with these small numbers, we could clearly differentiate between the three subpopulations.

The present study supports that the advanced DWI method using DWIC and FBA has a strong potential to improve our ability for early identification of heterogeneous imaging abnormalities underlying white matter injuries and disrupted maturation processes across different motor outcomes. Previous works<sup>25–28</sup> have investigated age-related white matter development in infants, mainly focusing on the voxel-wise measure of white matter maturation reporting the effects of myelination and brain water on increasing FA and decreasing mean diffusivity. The efficacy of neonatal DWIC analysis to investigate the developmental trajectory of whole-brain using different network topology (more clustered pair-wise connection) has been used at 45 weeks postconceptional age.<sup>29</sup> The topological locality of structural brain networks has been used to help predict neurobehavioral outcomes such as Bayley-III cognitive and motor scores for preterm infants (predictive correlation = 0.19 and 0.31 for cognitive and motor score).<sup>30</sup> Preterm infants when imaged at 38.6–47.1 weeks and utilizing FBA found a relationship between fixel-based measures (FD, FC, and FDC) with clinical risk factors in preterm, such as positive correlation with gestational age and negative correlation with days on requiring ventilation with FD, FC, and FDC.<sup>29</sup> This study also used a similar warping technique as ours. The warping of FOD may yield different results by causing a local shear and reduced number of fixels. Despite this limitation, when warping is used without bias, it could still be useful in the methodology in developing an objective biomarker. In a similar



**Table 3.** Mean and standard deviation of classification accuracy (Ac), sensitivity (Se), and specificity (Sp) obtained from the random forest algorithm of the multidimensional marker and patient metavariate.

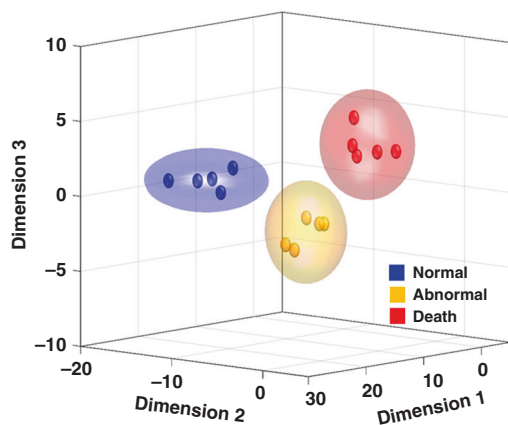
Modality	Marker	Without SMVCCA	With SMVCCA	With ranked SMVCCA	
DWIC	Log( $S(i,j)$ ) Ac	<b>0.894 ± 0.158</b> (19)	<b>0.923 ± 0.117</b> (4)	<b>0.987 ± 0.058</b> (5)	
	Se	1.000 ± 0.000 (19)	1.000 ± 0.000 (4)	0.997 ± 0.033 (5)	
	Sp	0.888 ± 0.222 (19)	0.942 ± 0.158 (4)	1.000 ± 0.000 (5)	
	FA		0.476 ± 0.132 (19)	0.570 ± 0.197 (3)	0.444 ± 0.146 (1)
			0.718 ± 0.215 (19)	0.748 ± 0.269 (3)	0.661 ± 0.279 (1)
			0.163 ± 0.239 (19)	0.423 ± 0.337 (3)	0.198 ± 0.223 (1)
	AD		0.411 ± 0.156 (19)	0.563 ± 0.158 (5)	0.437 ± 0.153 (3)
			0.523 ± 0.236 (19)	0.667 ± 0.271 (5)	0.606 ± 0.271 (3)
			0.119 ± 0.208 (19)	0.746 ± 0.322 (5)	0.283 ± 0.345 (3)
	RD		0.393 ± 0.172 (19)	0.483 ± 0.154 (5)	0.360 ± 0.153 (7)
			0.468 ± 0.275 (19)	0.532 ± 0.247 (5)	0.626 ± 0.337 (7)
			0.357 ± 0.369 (19)	0.605 ± 0.388 (5)	0.323 ± 0.356 (7)
	ADC		0.361 ± 0.147 (19)	0.477 ± 0.145 (6)	0.376 ± 0.166 (6)
			0.454 ± 0.271 (19)	0.535 ± 0.241 (6)	0.590 ± 0.364 (6)
			0.186 ± 0.259 (19)	0.607 ± 0.377 (6)	0.449 ± 0.441 (6)
FBA	FD		0.641 ± 0.172 (3)	0.628 ± 0.166 (3)	0.650 ± 0.174 (3)
			0.881 ± 0.184 (3)	0.844 ± 0.206 (3)	0.871 ± 0.205 (3)
			0.442 ± 0.362 (3)	0.436 ± 0.386 (3)	0.443 ± 0.388 (3)
	FC		0.327 ± 0.164 (3)	0.351 ± 0.150 (2)	0.300 ± 0.148 (3)
			0.447 ± 0.342 (3)	0.474 ± 0.355 (2)	0.263 ± 0.238 (3)
			0.599 ± 0.397 (3)	0.272 ± 0.327 (2)	0.510 ± 0.355 (3)
	FDC		0.601 ± 0.156 (3)	0.517 ± 0.153 (3)	0.477 ± 0.160 (3)
			0.853 ± 0.196 (3)	0.845 ± 0.202 (3)	0.843 ± 0.219 (3)
			0.404 ± 0.375 (3)	0.283 ± 0.339 (3)	0.223 ± 0.326 (3)
Clinical variable	Sex, GA, LOS		0.179 ± 0.112 (3)	0.183 ± 0.127 (2)	0.333 ± 0.156 (2)
			0.059 ± 0.127 (3)	0.040 ± 0.093 (2)	0.378 ± 0.284 (2)
			0.208 ± 0.298 (3)	0.130 ± 0.195 (2)	0.354 ± 0.358 (2)
Radiological variable	Intensity change and involvement on MRI		0.520 ± 0.143 (13)	0.160 ± 0.263 (3)	0.641 ± 0.153 (2)
			0.111 ± 0.274 (13)	0.147 ± 0.307 (3)	0.630 ± 0.315 (2)
			0.285 ± 0.284 (13)	0.082 ± 0.180 (3)	0.736 ± 0.280 (2)

() indicates the dimension providing the highest accuracy for each multidimensional marker; values in bold show that log( $S(i,j)$ ) has the highest accuracy value in each classification. Clinical and radiological variables of Table 1 were classified using the same classification algorithm for the comparison.

population as ours, using diffusion MRI at 6 months of age, there was only a trend to declining brain network integration and segregation with increasing neuromotor deficits following neonatal encephalopathy.<sup>31</sup> DTI and functional MRI using a passive motor task at 40 to 48 weeks' postconceptional age following perinatal brain injury showed FA and functional connectivity from the right supplemental motor area to be predictive of cerebral palsy at 2 years of age.<sup>32</sup> The Neonatal Research Network MRI pattern of neonatal brain injury was reported as a robust biomarker of neurodevelopmental outcome at 6–7 years of age.<sup>33</sup> A recent large cohort study<sup>34</sup> also reported that infants with better neurodevelopmental outcomes at the 1- and two-year follow-up showed higher FD, FC, and FDC in the corticospinal tract, midbrain, and corpus callosum, which suggests better information transfer capacity facilitated by an increased number of neurons, increased myelination, thicker bundles, and/or combinations.

Even though our study was retrospective in design, our MRI analysis was done in a blinded fashion to the groups. We are fully aware of the limitations of low statistical powered studies and the bias to overestimated effect sizes<sup>35,36</sup> and the need for more "n" for machine learning. We were not expecting the clear differentiation between the three populations with a difference between

bounds of 99% confidence intervals in Fig. 6. More studies and replication in more significant numbers of samples are needed to further establish whether these predictive prognostic markers will remain differentiated between our patient groups. Using Big Data approaches, we can now feed more data into this methodology. Also, another limitation is that despite the exceptional potential of clinical DWI data, it remains controversial whether current DWI tractography techniques can accurately reconstruct macroscopic structures of FOD functions and effectively remove false-positive tracts at the low angular acquisition of water diffusion.<sup>37,38</sup> Also, CSD and FBA are problematic with our DWI data with  $b$  value = 800 s/mm<sup>2</sup>. Ideally, the  $b$  value should be high (e.g., ~2500–3000 s/mm<sup>2</sup>) to reconstruct the FOD functions using CSD<sup>12</sup> and measure intra-axonal volume related to apparent FD.<sup>5,39</sup> When considering this practical problem, the proposed log( $S(i,j)$ ) marker might be limited in its ability to investigate the detailed mechanism about its biological origin. For instance, our preliminary data of abnormal motor and death groups showed a paradoxical increase in edge strength that may be related to the current pitfall of DWI tractography and our data quality, more likely tracking the wrong direction of the nearest fiber bundle at low spatial resolution. Nonetheless, we presumed that this spurious tracking would generate an exploratory marker that inevitably increases



**Fig. 6 Prediction of the eventual outcome made easier by ranked SMVCCA showing a figure plotting three of the five fused dimensions, which provided the most significant discrimination for three groups (Fig. 4a)** Each colored sphere indicates the three fused-dimensional feature of individual group subject. Each colored ellipse represents the upper limit of the three fused-dimensional features to predict each group at the confidence level of 99%.

false-positive tracts in constructing the edge of DWIC when the FOD functions of neighboring bundles have more spurious peaks with weak amplitudes as the ones from infants in the abnormal motor and death groups.

In conclusion, continued and systematic investigation using machine-learning techniques with clinical DWIC and FBA markers may improve the early prediction of neonatal motor outcomes. It may also allow identification of distinct patterns of white matter injuries, allowing more rapid and targeted intervention for improving long-term outcomes in term infants as a series of DWI-based studies<sup>40–42</sup> was consistently suggested to predict behavioral profiles, cognitive abilities, and language functions at 1–2 years old.

#### ACKNOWLEDGEMENTS

We thank the families who took part in the study, our colleagues in the Neonatal Intensive Care Unit, and the Developmental Assessment Clinic at Children's Hospital of Michigan. S.T. and J.-W.J. have been funded through NS114972 and NS117146 (S.T., P.I.) from NINDS, NIH.

#### AUTHOR CONTRIBUTIONS

J.-W.J.—study conception and design, data acquisition, data analysis and interpretation, statistical analysis, drafting manuscript and revision, final approval of the manuscript. M.-H.L.—data analysis and interpretation, statistical analysis, final approval of the manuscript. N.F.—drafting manuscript and revision, final approval of the manuscript. S. D.—drafting manuscript and revision, final approval of the manuscript. S.M.—drafting manuscript and revision, final approval of the manuscript. S.A.—manuscript revision, final approval of the manuscript. R.B.C.—manuscript revision, final approval of the manuscript. S.T.—study conception and design, data analysis and interpretation, drafting manuscript and revision, final approval of the manuscript.

#### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41390-021-01550-2>.

**Competing interests:** The authors declare no competing interests.

**Statement of consent:** The present study was approved by the University's Institutional Review Board, and a waiver of written informed consent was obtained to perform the analysis of existing data in the clinical archive.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### REFERENCES

- Tusor, N. et al. Prediction of neurodevelopmental outcome after hypoxic-ischemic encephalopathy treated with hypothermia by diffusion tensor imaging analyzed using tract-based spatial statistics. *Pediatr. Res.* **72**, 63–69 (2012).
- Martinez-Biarge, M. et al. Predicting motor outcome and death in term hypoxic-ischemic encephalopathy. *Neurology* **76**, 2055–2061 (2011).
- Brown, C. J. et al. Prediction of motor function in very preterm infants using connectome features and local synthetic instances. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (eds Nassir, N., Joachim, H., William, M. W., & Alejandro, F.) 69–76 (Springer, 2015).
- Jeong, J. W., Sundaram, S., Behen, M. E. & Chugani, H. T. Differentiation of speech delay and global developmental delay in children using DTI tractography-based connectome. *Am. J. Neuroradiol.* **37**, 1170–1177 (2016).
- Raffelt, D. A. et al. Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage* **144**, 58–73 (2017).
- Tournier, J. D., Calamante, F., Gadian, D. G. & Connelly, A. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage* **23**, 1176–1185 (2004).
- Wedeen, V. J. et al. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage* **41**, 1267–1277 (2008).
- Tournier, J. D., Calamante, F. & Connelly, A. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR Biomed.* **26**, 1775–1786 (2013).
- Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20**, 870–888 (2003).
- Alexander, D. C., Barker, G. J. & Arridge, S. R. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magn. Res. Med.* **48**, 331–340 (2002).
- Assemlal, H. E., Tschumperlé, D., Brun, L. & Siddiqi, K. Recent advances in diffusion MRI modeling: angular and radial reconstruction. *Med. Image Anal.* **15**, 369–396 (2011).
- Tournier, J. D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* **35**, 1459–1472 (2007).
- Smith, R. E., Tournier, J. D., Calamante, F. & Connelly, A. SIFT2: enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *Neuroimage* **119**, 338–351 (2015).
- Aydogan, D. B. & Shi, Y. Probabilistic tractography for topographically organized connectomes. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (eds Sebastien, O., Leo, J., Mert, R. S., Gozde, U. & William, W.) 201–209 (Springer, 2016).
- Shi, F. et al. Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS ONE* **6**, e18746 (2011).
- Zbyňek, Š. Rectangular confidence regions for the means of multivariate normal distributions. *J. Am. Stat. Assoc.* **62**, 626–633 (1967).
- Basser, P. J. & Pierpaoli, C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson.* **111**, 209–219 (1996).
- Wheeler-Kingshott, C. A. & Cercignani, M. About “axial” and “radial” diffusivities. *Magn. Reson. Med.* **61**, 1255–1260 (2009).
- Egorova, N. et al. Pervasive white matter fiber degeneration in ischemic stroke. *Stroke* **51**, 1507–13. (2020).
- Kelly, C. E. et al. Long-term development of white matter fibre density and morphology up to 13 years after preterm birth: a fixel-based analysis. *Neuroimage* **220**, 117068 (2020).
- Pannek, K. et al. Fixel-based analysis reveals alterations in brain microstructure and macrostructure of preterm-born infants at term equivalent age. *Neuroimage Clin.* **18**, 51–59 (2018).
- Pecheva, D. et al. Fixel-based analysis of the preterm brain: disentangling bundle-specific white matter microstructural and macrostructural changes in relation to clinical risk factors. *Neuroimage Clin.* **23**, 101820 (2019).
- Raffelt, D. A. et al. Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres. *Neuroimage* **117**, 40–55 (2015).
- Golugula, A. et al. Supervised regularized canonical correlation analysis: integrating histologic and proteomic measurements for predicting biochemical recurrence following prostate surgery. *BMC Bioinform.* **12**, 483 (2011).
- Lee, G. et al. Supervised multi-view canonical correlation analysis (SMVCCA): integrating histologic and proteomic features for predicting recurrent prostate cancer. *IEEE Trans. Med. Imaging* **34**, 284–297 (2015).
- Dudink, J., Kerr, J. L., Paterson, K. & Counsell, S. J. Connecting the developing preterm brain. *Early Hum. Dev.* **84**, 777–782 (2008).

27. Gao, W. et al. Temporal and spatial development of axonal maturation and myelination of white matter in the developing brain. *Am. J. Neuroradiol.* **30**, 290–296 (2009).
28. Gao, W. et al. Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc. Natl Acad. Sci. USA* **106**, 6790–6795 (2009).
29. Brown, C. J. et al. Structural network analysis of brain development in young preterm neonates. *Neuroimage* **101**, 667–680 (2014).
30. Kawahara, J. et al. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *Neuroimage* **146**, 1038–49. (2017).
31. Tymofiyeva, O. et al. Towards the “baby connectome”: mapping the structural connectivity of the newborn brain. *PLoS ONE* **7**, e31029 (2012).
32. Merhar, S. L. et al. Neonatal functional and structural connectivity are associated with cerebral palsy at two years of age. *Am. J. Perinatol.* **37**, 137–145 (2020).
33. Shankaran, S. et al. Neonatal magnetic resonance imaging pattern of brain injury as a biomarker of childhood outcomes following a trial of hypothermia for neonatal hypoxic-ischemic encephalopathy. *J. Pediatr.* **167**, 987–993 (2015).
34. Pannek, K. et al. Brain microstructure and morphology of very preterm-born infants at term equivalent age: associations with motor and cognitive outcomes at 1 and 2 years. *Neuroimage* **221**, 117163 (2020).
35. Button, K. S. et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**, 365–376 (2013).
36. Forstmeier, W., Wagenmakers, E. J. & Parker, T. H. Detecting and avoiding likely false-positive findings - a practical guide. *Biol. Rev. Camb. Philos. Soc.* **92**, 1941–1968 (2017).
37. Thomas, C. et al. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proc. Natl Acad. Sci. USA* **111**, 16574–16579 (2014).
38. Reveley, C. et al. Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. *Proc. Natl Acad. Sci. USA* **112**, E2820–E2828 (2015).
39. Genc, S. et al. Impact of b-value on estimates of apparent fibre density. *Hum. Brain Mapp.* **41**, 2583–2595 (2020).
40. Wee, C. Y. et al. Neonatal neural networks predict children behavioral profiles later in life. *Hum. Brain Mapp.* **38**, 1362–1373 (2017).
41. Girault, J. B. et al. White matter connectomes at birth accurately predict cognitive abilities at age 2. *Neuroimage* **192**, 145–155 (2019).
42. Ouyang, M. et al. Diffusion-MRI-based regional cortical microstructure at birth for predicting neurodevelopmental outcomes of 2-year-olds. *Elife* **9**, e58116 (2020).