ABSTRACT

BACKGROUND AND PURPOSE: Neonatal DTI enables quantitative assessment of microstructural brain properties. Although its use is increasing, it is not widely known that vast differences in tractography results can occur, depending on the diffusion tensor estimation methodology used. Current clinical work appears to be insufficiently focused on data quality and processing of neonatal DTI. To raise awareness about this important processing step, we investigated tractography reconstructions of the fornix with the use of several estimation techniques. We hypothesized that the method of tensor estimation significantly affects DTI tractography results.

MATERIALS AND METHODS: Twenty-eight DTI scans of infants born <29 weeks of gestation, acquired at 30-week postmenstrual age and without intracranial injury observed, were prospectively collected. Four diffusion tensor estimation methods were applied: 1) linear least squares; 2) weighted linear least squares; 3) nonlinear least squares, and 4) robust estimation of tensors by outlier rejection. Quality of DTI data and tractography results were evaluated for each method.

RESULTS: With nonlinear least squares and robust estimation of tensors by outlier rejection, significantly lower mean fractional anisotropy values were obtained than with linear least squares and weighted linear least squares. Visualized quality of tract reconstruction was significantly higher by use of robust estimation of tensors by outlier rejection and correlated with quality of DTI data.

CONCLUSIONS: Quality assessment and choice of processing methodology have considerable impact on neonatal DTI analysis. Dedicated acquisition, quality assessment, and advanced processing of neonatal DTI data must be ensured before performing clinical analyses, such as associating microstructural brain properties with patient outcome.

ABBREVIATIONS: LLS = linear least squares; WLLS = weighted linear least squares; NLLS = nonlinear least squares; RESTORE = robust estimation of tensors by outlier rejection; FA = fractional anisotropy

DTI enables in vivo assessment of WM microstructure and has become essential for quantification of brain abnormalities because it has been suggested to provide early biomarkers of neurodevelopment.1 Fiber tractography has the unique property to delineate specific WM pathways and is rapidly gaining in popularity because it may reveal substantial insights into disturbed brain connectivity and functionality of infants born preterm.2-4

There are many technical issues that may complicate the analysis of DTI data, including scanner type, hardware setup, acquisition parameters, and processing methodology.5,6 In addition, DTI applied in preterm infants is especially challenging because of specific clinical factors, such as the increased risk of subject motion, hemodynamic vulnerability, smaller head sizes, and higher heart and breathing rates compared with healthy adults. Therefore, before associations between tractography results and neurodevelopmental outcome can be established, it is of paramount importance that acquisition and processing of neonatal DTI data are performed with the highest standards possible.7,8 For example, different algorithms to estimate the diffusion tensor have been developed. These methods differ considerably in processing speed and dealing with data outliers. For instance, the linear least squares (LLS) method is widely used to estimate diffusion parameters but may lead to inaccuracy as it incorrectly assumes that data outliers are homogeneously distributed. Furthermore, there seems to be no consensus on how to practically define and handle data outliers. Awareness of these matters is essential because improper use may lead to inaccuracy; especially if data are compared when different estimation methods have been used. Unfortu-
nately, however, most studies using preterm brain DTI data have hardly focused on these important aspects, calling for a thorough investigation.

In the present study, trajectories of the fornix were reconstructed with fiber tractography for 28 preterm infants and compared when different diffusion tensor estimation approaches were used. Our hypothesis was that the chosen tensor estimation methodology significantly affects results of fiber tractography. This would demonstrate that an informed choice of diffusion tensor estimation is crucial for a reliable tractography analysis, which is especially relevant when artifact-sensitive DTI data of the preterm brain are involved.

**MATERIALS AND METHODS**

This study was approved by the institutional review board. Written informed parental consent was obtained for all subjects.

**Patients**

Between February 2011 and December 2012, preterm infants born before a gestational age of 29 weeks were recruited prospectively. MR imaging data were acquired at a postmenstrual age of 30 weeks (29 4/7 to 30 4/7 weeks). To avoid unnecessary data heterogeneity, infants with evidence of intracranial injury (intraventricular or cerebellar hemorrhage, WM abnormalities) observed with conventional MR imaging (see T1WI and T2WI protocols below) were excluded. Of the 217 eligible infants, 36 died before 30-week postmenstrual age; in 82 infants, the MR imaging scan could not be performed at 30-week postmenstrual age because of hemodynamic instability or logistic circumstances; and informed parental consent was not obtained for 20 infants. Of the remaining 79 infants, 36 had intracranial abnormalities, and 15 others were excluded from further analysis because different DTI acquisition settings were applied. This eventually resulted in 28 usable datasets.

**MR Imaging**

MR imaging procedures were carried out according to protocol: all infants were accompanied by trained staff only and were positioned in an MR imaging–compatible incubator (Lammers Medical Technology, Luebeck, Germany) that provided controlled temperature and humidity, MR-compatible pulse oximetry, and MR-compatible ventilation. Moldable earplugs and neonatal earmuffs protected the infants from auditory noise; no sedation was given.

Imaging data were acquired with a 1.5T Echoshoot scanner (GE Healthcare, Milwaukee, Wisconsin). Axial T2WI FSE was obtained with the following parameter settings: TR: 13,100 ms; TE: 139 ms; flip angle: 90°; section thickness: 1.2 mm; FOV: 190 × 190 mm²; acquisition matrix: 256 × 224; reconstruction matrix: 256 × 256 (voxel size: 0.74 × 0.74 × 1.23 mm³); acquisition time was 2:58 minutes. Axial 3D T1-spoiled gradient-echo MR imaging data were acquired by use of: TR: 9 ms; TE: 3 ms; flip angle: 15°; section thickness: 1.6 mm; FOV: 150 × 150 mm²; acquisition matrix: 224 × 224; reconstruction matrix 256 × 256 (voxel size: 0.59 × 0.59 × 1.6 mm³); acquisition time was 3:10 minutes. DTI was performed with the use of a single-shot EPI sequence with diffusion gradients in 25 noncollinear directions, TR: 11,725 ms; TE: 85.6 ms; section thickness: 3 mm; FOV: 220 × 220 mm²; acquisition matrix: 128 × 64; reconstruction matrix 256 × 256 (voxel size: 0.86 × 0.86 × 3 mm³); b value: 750 seconds/mm²; number of non–diffusion-weighted images: 3; acquisition time was 5:17 minutes.

**DTI Data Processing**

DTI data were analyzed with the use of ExploreDTI (http://www.exploredit.com) version 4.8.3. The diffusion-weighted images were first corrected for eddy currents, EPI distortion, and patient movement. The diffusion tensor was then estimated according to 4 different methods: 1) LLS; 2) weighted linear least squares (WLLS); 3) nonlinear least squares (NLLS), and 4) robust estimation of tensors by outlier rejection (RESTORE). Next, whole-brain tractography was performed for all datasets with the following parameters: fractional anisotropy (FA) threshold: 0.08; fiber length range: 15–500 mm; angle threshold 30°; and step size: 1 mm. Without loss of generality of this work, a single WM structure was investigated. Because of its important relation to cognition and high tracking reproducibility as the result of its unique shape, we performed tractography of the fornix. ROI placement was performed on the color-coded FA maps: 1) 1 “OR” ROI was placed in the axial plane at the level of the bilateral columns of the fornix, above the mammillary bodies; 2) 2 “AND” ROIs were placed in the coronal plane to encompass the corpus of the fornix and in the axial plane to include both crura of the fornix in the same section where the “OR” ROI was placed, and 3) 2 “NOT” ROIs were placed in the sagittal plane laterally to the seed region to exclude fibers from the anterior commissure. For each subject, tractography was repeated for each method of tensor estimation while the same subject-specific ROIs were used.

**Data Analysis**

Quality of the diffusion-weighted images was assessed with the outlier profiles of each dataset, after diffusion tensor estimation by use of RESTORE (Fig 2). The mean percentage of outliers per dataset was calculated by averaging the percentage artifacted voxels across the diffusion gradient orientations. In addition, tract parameters, including mean FA, mean diffusivity, mean fiber trajectory length (in mm), and number of fiber trajectories were computed for each dataset.

The overall quality of tractography was visually examined using the FA standard color map, and systematically evaluated (Table) by 2 authors independently. Both reviewers were blinded to the method of tensor estimation. The final score was the average of both total scores and ranged from 0–10 (Table). Statistical analysis was performed by use of SPSS version 20.0.0.1 (IBM, Armonk, New York). Intraclass correlations between both observers were calculated by use of a 2-way mixed model. Coefficients <0 were considered as no agreement; 0–0.20 as slight; 0.21–0.40 as fair; 0.41–0.60 as moderate; 0.61–0.80 as substantial; and 0.81–1 as almost perfect agreement. The Spearman correlation coefficient was used to investigate the correlation between the visualized quality of tract reconstruction and the mean outlier percentage per dataset. Independent-sample t tests served to test differences in visualized quality of tract reconstruction between data with <10% and >10% data outliers. Repeated-
measures ANOVA served to test differences in tract parameters between the diffusion tensor estimation methods, respectively. Difference in variability of tract parameters between the estimation techniques was tested with the Levene test for equality of variances. A value of $P < .05$ (2-sided) was considered statistically significant.

**RESULTS**

**Descriptive Statistics**

Twenty-eight infants (15 boys) were included in this study. Mean gestational age and birth weight were 27.7 weeks (SD: 1.1 weeks) and 1053 g (SD: 256 g), respectively. Mean postmenstrual age at image acquisition was 30.0 weeks (SD: 0.3 weeks). The mean percentage of outliers per dataset was 10.1% (SD: 1.3%).

**Outlier Evaluation**

Figures 3 and 4 are characteristic representations of data with poor and good quality of data, respectively. Interclass correlation between observers showed excellent agreement with high significance (intraclass correlation coefficient: 0.87; 95% confidence interval: 0.82–0.91; $P < .01$). Although there was some overlap among the methods of tensor estimation, visualized quality of reconstruction of the fornix was significantly higher with the use of the RESTORE algorithm, particularly in datasets with a high percentage (>10%, $n = 13$) of data outliers (Fig 5). Furthermore, visualized quality of tract reconstruction across all tensor estimation methods depended also on the presence of data artifacts because tract quality

---

**FIG 1.** Placement of ROIs. Tractography of the fornix was performed by placing 1 “OR” ROI (in blue), 2 “AND” ROIs (in green), and 2 “NOT” ROIs (in red) on color-coded fractional anisotropy maps.

**FIG 2.** Outlier profile of DTI data with high percentage of outliers.
was significantly correlated to the mean outlier percentage per dataset (Spearman correlation coefficient: \(-0.46; P < .01\)). This correlation was also tested for each tensor estimation method separately. The following Spearman coefficients were found: LLS: \(-0.48\) \((P = .01)\); WLLS: \(-0.47\) \((P = .01)\); NLLS: \(-0.57\) \((P < .01)\), and RESTORE: \(-0.36\) \((P = .06)\).

**Tract Parameters**

The impact of the diffusion tensor estimation method used on tract parameters is shown in Fig 6. There was a significant difference in mean FA value through the use of different diffusion estimation algorithms. Significantly lower mean FA values were obtained with NLLS and RESTORE than with LLS and WLLS. Furthermore, application of the RESTORE approach resulted in the lowest standard deviation of the mean FA value; LLS: 0.059; WLLS: 0.054; NLLS: 0.052; RESTORE: 0.051 (repeated-measures ANOVA, \(P < .05\)). Although not statistically significant, there was a trend toward an increased number of fiber trajectories in the following order of tensor estimation approaches: LLS, WLLS, NLLS, and RESTORE (Spearman correlation coefficient: 0.10; \(P = .32\)).

With the use of LLS, variability of mean FA values was significantly higher in datasets with more than 10% outliers compared with datasets with less than 10% outliers. With WLLS, NLLS, or RESTORE, there was no difference in variability of mean FA values with regard to the quality of the diffusion-weighted images (Fig 7).

**DISCUSSION**

This study emphasizes the paramount importance of quality assessment and dedicated use of processing methodology of neonatal DTI data before performing analysis. With our work, we demonstrated that 1) tract parameters are significantly affected by the chosen tensor estimation method and are estimated more reliably if data outliers are handled carefully; 2) robust estimation of the diffusion tensor results in significantly improved visualized quality of fiber reconstruction; 3) the mean percentage of data outliers of the diffusion-weighted images correlates significantly to visualized quality of tract reconstruction; and 4) data outliers are common and significantly affect subsequent DTI analysis if they are not taken into account.

Although the incidence of destructive types of brain injury with subsequent serious deficits is decreasing, preterm infants remain at considerable risk to develop cognitive and socio-emotional disabilities that persist into adolescence. Advanced MR imaging techniques, such as DTI, have already provided important valuable insights into WM microstructural properties of these “subtle” types of brain injury. Still, the neuropathologic correlates show high variation and inconsistency, and their workings are not completely understood. Moreover, preterm infants are vulnerable to respiratory and hemodynamic instability and movement artifacts, but, paradoxically, such work receives little attention in clinical research. This could be because of their emphasis on clinical results, or, more importantly, because of the lack of awareness that the choice of diffusion tensor estimation approach may affect the subsequent reconstruction of fiber pathways.

As shown in the present study, the choice of tensor estimation algorithms can significantly affect DTI tractography results, which may complicate the interpretation of specific findings. In this context, study populations can only be compared reliably when identical processing pipelines have been applied, necessitating...
the use of standardized guidelines before drawing conclusions with regard to outcome. Thus, strategies to limit image corruption should be incorporated into setups to acquire neonatal DTI data; 2) adjustment of parameter settings, by shortening diffusion time, applying stronger gradients, or by use of lower b-values; 3) oversampling gradient-sensitizing directions and removing corrupted diffusion-weighted images; and 4) applying more advanced tensor estimation methods. Because diffusion tensor estimation techniques differ considerably in principle, speed, and accuracy, awareness of the benefits and pitfalls is essential: the LLS method is fast and mostly used but assumes that errors are identically distributed, which can result in inaccurate estimation of the tensor. The WLLS method is slightly slower but provides more accurate results because it considers errors to be heterogeneously distributed. NLLS iteratively minimizes errors and results in more reliable estimation but needs considerably longer processing time and may get stuck in local optima during optimization. The RESTORE approach automatically detects and removes outliers before tensor estimation. This avoids manual and subjective identification of corrupted diffusion-weighted images and appears to be particularly valuable for data with frequent motion corruption. In summary, the reliability of DTI analyses is drastically improved when handling data outliers in an appropriate way. However, additional research is needed to determine what types of data processing can reliably be performed without affecting data quality. This report presses the need for careful data handling because corrupted data can significantly affect the final result. Although this will require longer processing time and perhaps the need to remove datasets completely, it will probably decrease the spread in the final analysis and therefore improve statistical significance and reduce sample size.

Limitations of this study are important and must be addressed. First, only datasets without evidence of injury were used, and this may have resulted in a selection bias. Although this policy provided a homogeneous study population, we did not investigate the impact of processing datasets with brain injury. Second, we applied an arbitrary boundary to define “good-quality” and “poor-quality” data: 10% data outliers. We used this threshold solely to illustrate the impact of poor data quality on DTI analysis; hence, we do not suggest that this 10% level should be used as a threshold for future studies to define poor data quality.
FIG 6. Impact of diffusion tensor estimation method on tract parameters. Tract parameters, such as fractional anisotropy (FA) (A), mean diffusivity (B), mean fiber trajectory length (C), and number of fiber trajectories (D) were affected by the tensor estimation method; mean FA value was significantly lower with use of the nonlinear least squares and robust estimation of tensors by outlier rejection techniques (paired sample t test, \( P < .05 \)).

FIG 7. Impact of data quality on variability of tract parameters. Diffusion-weighted images with high outlier percentages (>10%) resulted in a significantly increased variability of mean fractional anisotropy values compared with data with fewer data outliers (<10%) if linear least squares was used (Levene test for equality of variances, \( P < .05 \)). With application of robust estimation of tensors by outlier rejection, there was no difference in variability with regard to data quality.
CONCLUSIONS

As demonstrated with our tractography analysis of the fornix in the preterm brain, it is clear that the choice of diffusion tensor estimation methodology is crucial and that it has a considerable impact on subsequent analyses for studying microstructural brain properties. Given the insufficient attention in most clinical studies to date, this work raises the urgency to comply with the requirements to include state-of-the-art standardized research methodology wherever and whenever possible. Future studies should apply dedicated acquisition setups, standardized evaluation of data quality, and reliable processing of neonatal DTI data before performing analyses, such as associating microstructural brain properties with outcome.

Disclosures: Gabriel P. Krestin—UNRELATED: Consultancy: GEHC,* Comments: Strategic consultancy to GEHC Europe; Grants/Grants Pending: GEHC,* Siemens,* Bayer,* Bracco,* Philips,* *(money paid to institution).

REFERENCES

1. Ment LR, Hirtz D, Huppi PS. Imaging biomarkers of outcome in the developing preterm brain. Lancet Neurology 2009;8:1042–55
2. Bassi L, Ricci D, Volzono A, et al. Probabilistic diffusion tractography of the optic radiations and visual function in preterm infants at term equivalent age. Brain 2008;131:573–82
3. Ball G, Boardman JP, Aljabar P, et al. The influence of preterm birth on the developing thalamocortical connectome. Cortex 2013;49:1711–21
4. van der Aa NE, Leemans A, Northington FJ, et al. Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke? Stroke 2011;42:3410–14
5. Jones DK, Cercignani M. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR Biomedical 2010;23:803–20
6. Tournier JD, Mori S, Leemans A. Diffusion tensor imaging and beyond. Magn Reson Med 2011;65:1532–56
7. Heemskerk AM, Leemans A, Plaisier A, et al. Acquisition guidelines and quality assessment tools for analyzing neonatal diffusion tensor MRI data. AJNR Am J Neuroradiol 2013;34:1496–1505
8. Kozak LR, David S, Rudas G, et al. Investigating the need of triggering the acquisition for infant diffusion MRI: a quantitative study including bootstrap statistics. Neuroimage 2013;69:198–205
9. Plaisier A, Govaert P, Lequin MH, et al. Optimal timing of cerebral MRI in preterm infants to predict long-term neurodevelopmental outcome: a systematic review. AJNR Am J Neuroradiol 2014;35:841–47
10. Dyet LE, Kennea N, Counsell SJ, et al. Natural history of brain lesions in extremely preterm infants studied with serial magnetic resonance imaging from birth and neurodevelopmental assessment. Pediatrics 2006;118:536–48
11. Mukherjee P, Chung SW, Berman JL, et al. Diffusion tensor MR imaging and fiber tractography: technical considerations. AJNR Am J Neuroradiol 2008;29:843–52
12. Morris D, Nossin-Manor R, Taylor MJ, et al. Preterm neonatal diffusion processing using detection and replacement of outliers prior to resampling. Magn Reson Med 2011;66:92–101
13. Malamateniou C, Malik SJ, Counsell SJ, et al. Motion-compensation techniques in neonatal and fetal MR imaging. AJNR Am J Neuroradiol 2013;34:1124–36
14. Mathur AM, Neil JJ, McKinstry RC, et al. Transport, monitoring, and successful brain MR imaging in unsedated neonates. Pediatr Radiol 2008;38:260–64
15. Veraart J, Rajan J, Peeters RR, et al. Comprehensive framework for accurate diffusion MRI parameter estimation. Magn Reson Med 2013;70:972–84
16. Koay CG, Chang LC, Carew JD, et al. A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging. J Magn Reson Med 2006;56:115–25
17. Salvador R, Pena A, Menon DK, et al. Formal characterization and extension of the linearized diffusion tensor model. Hum Brain Mapp 2005;24:144–55