

Contents lists available at ScienceDirect

NeuroImage



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

Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data



V. Sairanen^{a,b,*}, A. Leemans^c, C.M.W. Tax^d

^a Department of Physics, University of Helsinki, Helsinki, Finland

^b HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

^c Image Sciences Institute, University Medical Center Utrecht, Utrecht, Netherlands

^d Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, United Kingdom

ABSTRACT

The accurate characterization of the diffusion process in tissue using diffusion MRI is greatly challenged by the presence of artefacts. Subject motion causes not only spatial misalignments between diffusion weighted images, but often also slicewise signal intensity errors. Voxelwise robust model estimation is commonly used to exclude intensity errors as outliers. Slicewise outliers, however, become distributed over multiple adjacent slices after image registration and transformation. This challenges outlier detection with voxelwise procedures due to partial volume effects. Detecting the outlier slices before any transformations are applied to diffusion weighted images is therefore required. In this work, we present i) an automated tool coined SOLID for slicewise outlier detection prior to geometrical image transformation, and ii) a framework to naturally interpret data uncertainty information from SOLID and include it as such in model estimators. SOLID uses a straightforward intensity metric, is independent of the choice of the diffusion MRI model, and can handle datasets with a few or irregularly distributed gradient directions. The SOLID-informed estimation framework prevents the need to completely reject diffusion weighted images or individual voxel measurements by downweighting measurements, SOLID detects outliers with a high sensitivity and specificity, and can achieve higher or at least similar sensitivity and specificity compared to other tools that are based on more complex and time-consuming procedures for the scenarios investigated. SOLID was further validated on data from 54 neonatal subjects which were visually inspected for outlier slices with the interactive tool developed as part of this study, showing its potential to quickly highlight problematic volumes and slices in large population studies. The informed model estimation framework was evaluated both in simulations and in vivo human data.

1. Introduction

Diffusion MRI (dMRI) is sensitized to the microscopic motion of particles (Basser et al., 1994; Stejskal and Tanner, 1965) and is extensively being used to study brain connectivity and tissue microstructure in case of normal development, disorders, and training (Jones, 2008; Langen et al., 2012; Odish et al., 2015; Baum and Stevenson, 2016; Hämäläinen et al., 2017). Diffusion weighted images (DWI) are, however, affected by a range of artefacts, including image misalignment due to subject motion and geometrical distortions due to field inhomogeneities and eddy currents (Tournier et al., 2011; Tax et al., 2016; Andersson and Skare, 2010; Pierpaoli, 2010; Heemskerk et al., 2013; Kennis et al., 2016). Commonly used correction strategies for these artefacts are based on the post-acquisition registration and geometrical transformation of the images, and are integrated into various software tools for dMRI processing (Pierpaoli and Walker, 2010; Jezzard et al., 1998; Mangin et al., 2002; Nielsen et al., 2004; Rohde et al., 2004; Leemans et al., 2009).

In addition to geometrical artefacts, intensity errors that are nonrelated to the microscopic motion of particles can impede an unequivocal characterization of the true physical diffusion process (e.g. Vos et al., 2017). For example, subject head motion during the acquisition of a DWI slice can induce artificial signal decreases whereas hardware issues could also lead to erroneous signal increases (Jones and Cercignani, 2010; Le Bihan et al., 2006) which cannot be ameliorated with geometrical transformations, even though the error sources for both artefacts could be the same. This is especially harmful for clinical datasets with relatively few directions per shell as artefactual measurements can have high effect on the final results. Although multiple tools have been developed to detect intensity errors, the importance and complexity of this problem has recently been re-emphasized (Andersson et al., 2016). Voxelwise detection strategies (Cook et al., 2006; Lucas et al., 2010; Pierpaoli and

https://doi.org/10.1016/j.neuroimage.2018.07.003

Received 23 October 2017; Received in revised form 22 May 2018; Accepted 2 July 2018 Available online 5 July 2018 1053-8119/© 2018 Elsevier Inc. All rights reserved.

^{*} Corresponding author. HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, POB 340, Haartmaninkatu 4, 00029, Helsinki, Finland.

E-mail address: viljami.sairanen@helsinki.fi (V. Sairanen).

Walker, 2010; Pannek et al., 2012; Chang et al. 2005, 2012; Collier et al., 2015; Tax et al., 2015; Mangin et al., 2002) based on robust estimators are typically performed after correction for geometric distortions, but the geometric transformations cause slicewise intensity artefacts to spread over multiple locations. Slicewise intensity errors, in particular, become distributed over multiple adjacent slices and appear as gradual intensity bands within the resampled DWIs. The edge regions of such gradual intensity bands are particularly problematic for these voxelwise robust estimators due to partial volume effects; the local intensity is an interpolation between normal and outlier data due to the transformation and resampling, and as a result such partial volume outliers could fall within the normal noise variation. Slicewise intensity outliers must therefore ideally be detected prior to any DWI geometric transformation (Morris et al., 2011; Andersson and Sotiropoulos, 2016; Marami et al., 2016; Marami et al. 2016, 2016), which is not generally adopted (e.g. Lauzon et al., 2013). In addition, using voxelwise strategies to detect slicewise outliers depends on the choice of voxelwise diffusion signal model or representation, and does not fully exploit the information present in the signals of the rest of the slice.

Several strategies have been proposed specifically for slicewise outlier detection (Zhou et al., 2011; Lauzon et al., 2013; Liu et al., 2010; Oguz et al., 2014; Jiang et al., 2006; Li et al., 2013; Andersson et al. 2016, 2017; Andersson and Sotiropoulos, 2016; Pannek et al., 2015; Scelfo et al., 2012), which are either based on the slicewise comparison of intensity metrics within the same DWI or on using slicewise information across DWIs. The algorithm of Liu et al. (2010); Oguz et al., 2014 (as implemented in DTI-Prep) does within-DWI detection based on the normalized cross-correlation of adjacent slices and rejects the whole DWI volume if it contains multiple outliers. However, an affected DWI may still contain useful information, and ideally only the corrupted slices should be detected and their effect on subsequent analysis be reduced. Jiang et al. (2006), Li et al. (2013) and Marami et al. (2016a,b) based their detection to a morphological closing operation (implemented in DTI Studio) to calculate voxelwise indices that are used to guide the model estimation. Both of the aforementioned algorithms can result in false positives in the case of multiple adjacent or interleaved outliers. Zhou et al., 2011 used a computer vision approach called "local binary patterns" that describe pixel neighborhoods in an image (Ojala et al., 1996) to detect slicewise outliers across DWIs, whereas Lauzon et al., 2013 based their detection method to the sum of squared tensor model residuals for each slice. Andersson et al., 2016 proposed the use of a non-parametric Gaussian process (Andersson and Sotiropoulos, 2015) to model and predict diffusion-weighted signals based on the measured DWIs and integrate this with subject motion and distortion corrections (implemented in FSL EDDY). If the measured DWI slice intensities significantly differ from their predicted counterparts, the slice is deemed an outlier and can be replaced by its prediction in further analysis. While this integrated framework provides an elegant solution to the detection of slice outliers before geometrical transformation, the signal prediction sets constraints on the minimal data acquisition and can be time consuming.

In this work, we sought to develop a fast and less complex approach that lifts these restrictions and thus would be compatible with datasets acquired using less gradient directions. To achieve this, we present i) an automated tool coined SOLID (Slicewise OutLIer Detection) to identify slicewise outliers across DWIs before geometrical transformations are applied and ii) a framework to minimize outlier slice impact on estimations which is demonstrated with a tensor model. SOLID is based on a straightforward intensity metric and does not rely on the time-consuming model estimation and predictions of the dMRI signal. Its sensitivity and specificity are extensively evaluated on simulated DWIs with a range of clinical b-values, SNRs, and misalignments of the input images, as well as on 156 data sets from 54 different neonatal subjects. In addition, SOLID is compared to three previously published tools for slicewise outlier detection: DTIPrep (Liu et al., 2010; Oguz et al., 2014), DTI Studio (Jiang et al., 2006; Li et al., 2013), and FSL's EDDY (Andersson et al. 2016, 2017; Andersson and Sotiropoulos, 2016). Instead of removing the entire DWI or

replacing slicewise outliers with predictions, we propose here to incorporate the measurement 'uncertainty' derived from SOLID into the estimation (Knutsson and Westin, 1993; Tax et al., 2017; Li et al., 2013). Whereas previous work has minimized the effect of outliers by recuding their weight based on their voxelwise model residuals (Mangin et al., 2001; Meer et al., 1991), here the reduction is based on the combined effect of the brain voxel intensities within the outlier slice without the voxelwise modelling.

2. Methods

The SOLID framework consists of two parts: i) slicewise outlier detection that is based on differences in DWI intensity histograms between slices across DWIs and ii) model estimation informed by the detection results. Fig. 1 gives a schematic overview of the different steps of the SOLID framework. First, we revise the statistical methods of outlier detection based on the Z-score and its robust modified counterpart (Iglewicz and Hoaglin, 1993; Norman and Streiner, 2007) and explain how this was implemented in SOLID. Second, we detail how the information from SOLID can be integrated into a model estimation framework. Third, we present the simulation experiments and real data acquisitions for the evaluation of the detection and model estimation, and finally provide methodological details of the tools that are evaluated for comparison.

2.1. Slicewise OutLIer Detection: SOLID

To examine the likelihood of a slice k of DWI l being an outlier, it is useful to design a slicewise summary feature or 'observation' $y_{k,l}$ that can be compared across slices and/or DWIs. The Z-score can subsequently be used as a statistical approach to identify outliers from the multiple observations y of a random variable Y which are drawn from a normal distribution with a mean μ and a standard deviation σ i.e. $Y \in N(\mu, \sigma^2)$. The steps of SOLID are as follows (Fig. 1A–C, note that this is on the raw non-registered data):

- A) The intensity histograms of normal and outlier slices have distinct characteristics (Scelfo et al., 2012).
- B) The summary statistic, or the 'observation' y, is a slicewise intensity metric calculated within a brain mask for each slice k and each DWI l per shell, denoted by $y_{k,l}$. In the remainder of the manuscript, we use the variance of the intensities as the metric, but other choices are also possible and are further elaborated upon in the Discussion.
- C) Using the Z-score, an observation can be deemed an outlier if the difference between the measurement and the sample mean \overline{y} divided by the sample standard deviation *s* is large. More specifically, in the case of slicewise outlier detection, the mean \overline{y}_k and the standard deviation s_k can be calculated for each slice *k* across DWIs per shell, resulting in a slice- and DWI-specific Z-score $z_{k,l}$ (eq. (1)):

$$z_{k,l} = \frac{y_{k,l} - \overline{y}_k}{s_k}.$$

A challenge when using the mean and standard deviation of observations for outlier detection is that in the presence of multiple outliers, a masking effect might occur, and milder outliers could remain undetected. The modified Z-score ζ (Iglewicz and Hoaglin, 1993) was introduced to overcome this drawback by replacing the mean \overline{y} with the median \tilde{y} and the standard deviation *s* with the median absolute deviation (MAD) (Mangin et al., 2002). The modified Z-score $\zeta_{k,l}$ is calculated with the help of the slicewise median \tilde{y}_k and MAD_k:

$$\zeta_{k,l} = \frac{\mathbf{y}_{k,l} - \tilde{\mathbf{y}}_k}{\mathbf{MAD}_k},\tag{2}$$



Fig. 1. SOLID framework: the detection of slicewise outliers from DWIs and the informed model estimation. A-B) An intensity metric *y* (here variance) is computed across all voxels **x** for each slice *k* in DWI *l*, denoted by $y_{k,l}$. C) The modified Z-score $\zeta_{k,l}$ is calculated from $y_{k,l}$, the slicewise median \tilde{y}_k and the median absolute deviation MAD_k. The color scale already reveals some outlier slices with high values. D) Geometrical misalignment correction step gives image transformation matrices T_l for each volume. Red dashed lines visualize the axial plane before and after transformation relative to the signal decrease slice outlier. E1) Voxelwise interpolated modified Z-score s $\xi_l(\mathbf{x})$ are obtained using the same transformations T_l for each modified Z-score volume $\xi_l(\mathbf{x})$. E2) A voxelwise SOLID weight $S_l(\mathbf{x}')$ is linearly interpolated from the $\xi_l(\mathbf{x}')$ to obtain the certainties of data points. E3) The SOLID-informed model estimation.

$$MAD_k = 1.4826 \cdot \text{median}(|y_{k,l} - \tilde{y}_k|).$$
(3)

The slicewise modified Z-score map (Fig. 1C) provides a quick but very useful overview of the data: suspicious slices with a high modified Zscore can be readily displayed for visual inspection. To do actual outlier 'detection and rejection' to reduce the impact on model fitting, one could straightforwardly set an arbitrary threshold on the modified Z-scores and subsequently discard the DWI intensities with the modified Z-score exceeding this threshold during estimation. However, the geometric transformation step to correct for motion and distortion prior to model estimation results in gradual intensity bands (Fig. 1D), which challenges the choice of an appropriate threshold setting because intensities in voxels that are only partially affected by outliers could completely be rejected depending on the threshold setting. Even though the SOLID output allows the user to detect and reject outlier slices in this way, we propose not to adopt this approach but instead to downweight data points based on the degree of partial voluming with the outlier data. This SOLID-based estimation after geometric transformation is described in the next section.

2.2. SOLID-informed estimation after geometric transformation

Incorporating the degree to which a voxel is affected by an outlier measurement to the model estimation is in line with the signal/certainty philosophy, i.e. naturally separating the data into a signal and certainty part (Knutsson and Westin, 1993). Here, we propose to derive the estimate for 'certainty' from interpolated modified Z-scores. We will refer to this certainty as 'SOLID-weight'. The steps of the SOLID-based estimation are as follows (Fig. 1E, note that geometric transformation can be done with any preferred method that can output the transformation for each DWI l (T_l), even in a slice-to-volume manner (Ferrante and Paragios, 2017)):

E1) Transforming the slicewise modified Z-scores $\zeta_{k,l}$ into voxelwise SOLID weights after geometric transformation requires first to map them to 3D volumes ξ_l (l = 1, ..., n with n being the number of

2

DWIs) so that all the voxels in the slice k of volume l have the same value, i.e. $\xi_l(i,j,k) = \zeta(k,l)$. Here, i and j denote the spatial coordinates in the slice-plane. Note that the 3D volumes $\xi_l(i,j,k)$ contain the same information as the modified Z-score map $\zeta(k,l)$, but have the same dimensions as the image data. To simplify the notation, the 3D volumes can be written as $\xi_l(\mathbf{x})$ with \mathbf{x} being the voxel coordinate. Subsequently, the same volumetric transformations T_l that are used to correct DWIs for geometric distortions are applied to the 3D volumes $\xi_l(\mathbf{x})$, resulting in voxelwise transformed and interpolated modified Z-score maps $\xi'_l(\mathbf{x}') = \xi_l(T_l(\mathbf{x}))$.

E2) Voxelwise SOLID weights $S_l(\mathbf{x})$ in the transformed image space are derived from the modified Z-scores $\xi_l(\mathbf{x})$ by scaling them linearly between 0 (outlier) and 1 (reliable data point) using manually chosen lower and upper thresholds t_L and t_U as shown in eq. (4). In contrast to the voxelwise Geman-McClure weights (Meer et al., 1991; Mangin et al., 2002), SOLID weights are based on the data from all the brain voxels within a slice thus providing additional statistical power for the outlier detection, and do not depend on the residuals of an estimated model.

$$S_{l}(\mathbf{x}') = \begin{cases} 0 & , \text{ if } \xi_{l}'(\mathbf{x}') > t_{U} \\ \frac{\xi_{l}'(\mathbf{x}') - t_{U}}{t_{L} - t_{U}} & , \text{ if } t_{L} \le \xi_{l}'(\mathbf{x}') \le t_{U} \\ 1 & , \text{ if } \xi_{l}'(\mathbf{x}') < t_{L} \end{cases}$$

$$\tag{4}$$

Thresholds $t_U > t_L \ge 0$ can be tuned to adjust for the normal variation of the modified Z-scores within the data. When variance is chosen as the slicewise intensity metric $y_{k,l}$, we recommend a lower threshold of $t_L =$ 3.5 based on our experiments and the previous literature (Iglewicz and Hoaglin, 1993). Instead of normalizing to the maximum modified Z-score in $\xi'_l(\mathbf{x}')$ to obtain a SOLID weight between 0 and 1, we propose the use of an upper threshold $t_U \le \max(\xi'_l(\mathbf{x}'))$ beyond which the SOLID weight is clipped to 1. We found a value two to three times larger than the lower threshold (by default $t_U = 10$) appropriate to downweight outlier data points, while preventing extreme outliers with large modified Z-score from dominating the process. Practically, this means that $\xi'_l(\mathbf{x}') < t_L$ will have a weight of 1, $\xi'_l(\mathbf{x}') > t_U$ will have a weight of 0, and the remaining $\xi'_l(\mathbf{x}')$ will have a value between 0 and 1 during the model estimation; i.e. the edge regions of gradual intensity bands are naturally downweighted based on the degree of partial voluming with the outlier slice.

E3) The SOLID weights are subsequently used to inform a model estimator, where different approaches can be taken depending on the model and the estimator. For the iteratively re-weighted linear least squares (IWLLS) estimator (Veraart et al., 2013), a straightforward incorporation is to multiply the diagonal of the weight matrix by the corresponding SOLID weights $S(\mathbf{x})$ combined from all shells $S_l(\mathbf{x})$ during each iteration step *m* giving weights estimates $\tilde{\boldsymbol{W}}_m(\mathbf{x})$ for each voxel \mathbf{x} based on the design matrix \boldsymbol{X} and voxelwise tensor element estimates $\hat{\boldsymbol{\beta}}(\mathbf{x}')$:

$$\tilde{\boldsymbol{W}}_{m}(\mathbf{x}') = \operatorname{diag}(S(\mathbf{x}') \cdot \exp(2\boldsymbol{X}\widehat{\boldsymbol{\beta}}(\mathbf{x}')_{m-1})).$$
(5)

2.3. Simulation experiments

Simulations were based on a part of a human full-brain dataset of the MASSIVE database (Froeling et al., 2017). The data were acquired on a single subject in eight sessions and consisted of 170 non-diffusion weighted images and 250, 500, and 500 images with b-values of 500 s/mm^2 , 1000 s/mm^2 , and 2000s/mm^2 , respectively. A ground truth (GT) dataset was obtained by anisotropically smoothing the data (full width at half maximum of 6 mm) (Leemans et al., 2009) and

subsequently estimating diffusion tensor (DT) and kurtosis tensor (KT) using the IWLLS estimator in ExploreDTI (Leemans et al., 2009; Veraart et al., 2011). Finally, the estimated tensors were used to simulate the GT data with non-diffusion-weighted signal and diffusion signal on two shells with b-values of 1000 s/mm^2 and 2000 s/mm^2 . The number of diffusion gradient directions per shell is detailed in Table 1.

2.3.1. Outlier detection

To evaluate the sensitivity and specificity of SOLID, receiver operating characteristic (ROC) curves were generated by setting hypothetical thresholds on the modified Z-score values $\zeta_{k,l}$ and counting the true/false positive/negative slices. A similar approach was used to evaluate positive predictive value through precision - recall curves (PRC). Note that this threshold is 'hypothetical' because in the data/certainty approach, a positive finding does not directly lead to the exclusion of the measurement. In all outlier detection simulations, each shell consisted of the same 30 diffusion gradient orientations (Jones et al., 1999). Outlier slices were introduced in DWIs in different setups detailed in Table 1. Positions of the outlier slices were fixed across all affected DWIs to maximize the effect. For each setup, both a complete signal loss (-100%) and a modest signal increase (+50%) artefacts were generated using four SNR levels (8, 16, 32, and infinite) of Rician noise. The random selection of artefactual DWIs and slicewise outliers was repeated to form 1000 unique sets of 30 DWIs for each experiment.

To study whether initial geometrical misalignments of structures resulting from subject motion were problematic for the slicewise comparison in SOLID, 5-degree rotations around the left-right and the anterior-posterior axes were applied to the artefactual DWIs and the experiments described above were repeated. A common brain mask calculated as a union of all DWI brain masks was used in outlier detection simulations.

2.3.2. Comparison with existing algorithms for slicewise outlier detection

SOLID is compared to three previously published tools for slicewise outlier detection: DTIPrep (Liu et al., 2010; Oguz et al., 2014), DTI Studio (Jiang et al., 2006; Li et al., 2013), and FSL's EDDY (Andersson et al., 2016; Andersson and Sotiropoulos, 2016). For the former two tools, we qualitatively compare the results on a simulated dataset whereas for FSL EDDY we did a more complete ROC and PRC profile-based comparison.

DTIPrep (Liu et al., 2010) uses a voxelwise normalized cross-correlation metric between successive slices within each DWI. The algorithm assumes that normalized cross-correlation values calculated at the same slice pair positions across DWIs follow normal distributions. If the normalized cross-correlation score of two successive DWI slices deviates more than 3.5 standard deviations from the mean normalized cross-correlation score at a given slice position, the pair is considered to have an outlier slice. This method by design does not identify which one of the adjacent slices was the outlier. To circumvent this problem, DTI-Prep rejects the whole DWI volume if it contains a user-specified number of outlier slices.

DTI Studio (Li et al., 2013) performs a morphological closing operation perpendicular to the acquisition plane and subtracts the result from the original DWI. If there is a slicewise signal decrease outlier, it will get 'closed', and the subtraction image shows bright voxels in that area. DTI Studio takes a similar data/certainty approach as proposed here, by estimating the certainty with "corrected inter-slice intensity discontinuity" metric and using it as a factor in model estimation weights.

FSL's EDDY (Andersson and Sotiropoulos, 2016; Andersson et al., 2017) algorithm makes slicewise predictions based on a Gaussian process and compares these predictions with the acquired slices. The detection criterion assumes that voxelwise intensity differences between measured and predicted slices follow a normal distribution. Slices with a difference larger than a predefined number of standard deviations from the mean, are considered outliers. We performed comparison for the ROC curve and the PRC profile by varying the standard deviation threshold setting for

Table 1

The overview of the performed simulations and real data experiments.

Section/ Figs. [suppl.]	Dataset properties							
	Туре	b-value [s/ mm ²]	SNR	Outlier intensity deviations [%]	Number of outliers per DWI	Number of artefactual DWIs	Rotation(s) [°]	Number of random repetitions
	Simulations: slicewise outlier detection							
3.1.1/ 2A-B [S1A- B]	Full-brain	1000	8,16,32,∞	-100, 50	1	1/30	n/a	1000
3.1.1/ 2C-D [S1C- D]	Full-brain	1000	8,16,32,∞	-100, 50	5	8/30	n/a	1000
3.1.1/ 2E-F [S1E- F]	Full-brain	2000	8,16,32,∞	-100, 50	5	8/30	n/a	1000
3.1.1/[S2]	Full-brain	1000	8,16,32,∞	-50, -10	5	8/30	n/a	1000
3.1.1/ 3 [S3]	Full-brain	1000	8,16,32,∞	-100, 50	5	8/30	5 AP, LR	1000
3.1.2/ 4B-D	Full-brain	1000	16	-100, 50	16	1/30	n/a	1
3.1.2/4E [S4]	Full-brain	1000	16	-100, 50	5	8/30	5 LR	1000
	Simulations: informed model estimation							
3.1.3/[\$5]	Gray matter	1000, 2000	32	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/5	White matter	1000, 2000	32	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/[S6]	Gray matter	1000, 2000	16	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/[\$7]	White matter	1000, 2000	16	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/[\$8]	Gray matter	1000, 2000	64	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/[\$9]	White matter	1000, 2000	64	-100 to 50	n/a	0 to 16/60	n/a	1000
3.1.3/ 6	White matter	1000	16	-100	2	2/30	5 LR	1
3.1.3/[\$10]	White matter	1000	16	-100, 50	5	8/30	5 LR	1000
	Human data experiments							
3.1.4/7	Full-brain	700	Clinical data of 156 sets of 15 DWIs with outliers, signal deviations, and subject motions.					
3.1.4/7	Full-brain	1000	Co-operative subject data of 64 DWIs with negligible motion or artefacts.					

EDDY from 0 to 300 and the modified Z-score threshold for SOLID from 0 to 300. The test data consisted of 30 DWIs with $b = 1000 \text{ s/mm}^2$ of which eight were randomly replaced with artefactual volumes that had five randomly located outlier slices with either a 100% signal decrease or a 50% signal increase and 5-degree rotation around left-right axis. The selection of artefactual DWIs, outlier slices, and the addition of Rician noise with SNR 16 was repeated 1000 times.

2.3.3. SOLID-informed estimation: the effect of introducing uncertainty

The small degrees of partial voluming with outlier slices or small signal deviations do not significantly affect model estimates. Accompanying the measurement with a certainty (SOLID weight) allows for reducing the effect of different degrees of signal deviation. To evaluate the effect of the degree of signal deviation and the downweighting on diffusion measure estimates, voxels in the corpus callosum (CC) and in the deep gray matter (GM) were selected from the GT to represent typical tissue signals. The CC voxel had fractional anisotropy (FA), mean diffusivity (MD), mean kurtosis (MK), kurtosis anisotropy (KA), and radial kurtosis (RK) (Poot et al., 2010) values of 0.85, $0.87 \cdot 10^{-3} \text{ mm}^2/\text{s}$, 1.5, $0.94 \cdot 10^{-3} \text{ mm}^2/\text{s}$, 0.75, 0.10, and 0.85, respectively. Diffusion signals were simulated on two shells with a 60 gradient direction scheme (Jones et al., 1999) with diffusion weightings of 1000 s/mm² and 2000 s/mm² along with five non-diffusion weighted signals.

Diffusion signals were simulated with different diffusion weightings detailed in Table 1. Multiple setups were generated by randomly selecting a varying amount of directions (up to 16 out of 60) and introducing signal deviations from a complete signal loss (-100%) to a signal increase (+50%) with 10% steps. This setup encompasses situations where outliers with different signal deviations are fully or partially interpolated into the result. The degree of the signal deviation in each

setup was the same for all the directions. For each setup, 1000 Rician noise iterations were introduced with different SNRs, and the DKI equation was fitted using IWLLS (Veraart et al., 2013). The certainty (SOLID weight) was gradually reduced for the outlier signals from 100% to 0% with 10% steps (eq. 5).

Finally, we compare the performance of SOLID-informed IWLLS tensor estimation with normal IWLLS tensor estimation and REKINDLE (Tax et al., 2015) in gradual intensity band regions.

2.4. Real data experiments

The usage of neonatal data was approved by the Ethics Committee of the Helsinki University Hospital and the usage of the adult data was approved by the University of Helsinki Ethical Review Board of Humanities and Social and Behavioural Sciences. All subjects or parents of the newborn subjects had given their written consent.

The sensitivity and specificity of SOLID outlier detection was evaluated with the data of 54 neonatal subjects scanned as a part of a previous study (Stjerna et al., 2015). Full-brain axial DWIs were acquired using Philips Achieva 1.5T scanner (Best, The Netherlands) with an 8-channel phased array head coil. A single-shot echo planar imaging sequence with an acquisition matrix of 128×128 , voxel size of $1.75 \times 1.75 \times 2.0 \text{ mm}^3$ and TR/TE of 6700/58 ms was used to acquire 15 DWIs based on over-plus gradient scheme with *b*-value of 700 s/mm^2 and one non-diffusion-weighted image. The acquisition was repeated up to four times depending how co-operative the subject was, resulting in 156 sets of DWIs which were analyzed separately.

All DWIs were visually inspected for slicewise artifacts in coronal and sagittal views. If any axial outlier slices were identified, they were confirmed in the axial view. Outliers based on this visual detection were used as the gold standard for the original validation of SOLID. As we observed multiple false positives in form of a low positive predictive value we revised the validation with a subsequent visual analysis of outliers found by SOLID to confirm findings. This revealed that SOLID was capable of identifying more subtle outliers that human observer had missed in the original visual detection. The true and false positive findings were updated based on the revised inspection. One subject with severe motion artefacts was studied in more detail to visualize how SOLID weights are affected by different lower threshold settings t_L (eq. (4)).

In addition, an adult subject was used to visualize SOLID performance on data without any visible artefacts and to highlight how the lower threshold setting $t_L > 0$ can prevent unnecessary downweighting of good quality data. The full-brain DWI acquisition of the subject was done as a part of previous study (Hämäläinen et al., 2017) using Siemens Skyra 3T scanner (Erlangen, Germany) with a 32-channel head matrix coil. The acquisition matrix was 120×120 with an isotropic voxel size of $2 \times 2 \times 2$ mm³ and TR/TE of 9600/81 ms. Five non-diffusion weighted images and DWIs with diffusion sensitizing gradients in 64 directions provided by the vendor with a *b*-value of 1000 s/mm² were acquired.

3. Results

First, we will show simulation results related to the outlier detection procedure, i.e. the sensitivity and specificity of SOLID along with a comparison to other existing outlier detection tools. Second, we show the effect of outliers on tensor derived measures and how this effect can be reduced by applying the SOLID weight. Third, we show the evaluation of SOLID on the neonatal dataset. Finally, we highlight two example subject cases where the modified Z-score included as data certainty improves the results. A summary of all experiments is given in Table 1.

3.1. Simulation experiments

3.1.1. Outlier detection

Fig. 2 shows the ROC curves of SOLID in different outlier simulation settings. For each simulation setup, an Area-Under-Curve (AUC) value and the optimal modified Z-score $\xi_l(\mathbf{x})$ threshold is reported. The optimal threshold was calculated as the minimum Euclidean distance from the ROC curve to the 100% sensitivity and specificity, i.e., the top left of the plot. The ROC curve was calculated by applying modified Z-score thresholds from $\xi_l(\mathbf{x}) = 0$ i.e. 100% sensitivity and 0% specificity to $\xi_l(\mathbf{x}) = 10$ to cover the range of interesting values including the optimal modified Z-scores. Respective results for PRC profiles are reported in Supplementary Fig. 1.

Fig. 2A and B) displays the simulation results of 'a typical DTI dataset'. In this setup, SOLID performed well over all tested SNR levels with the lowest ROC AUC value being approximately 0.98 with SNR 8. C-D) demonstrates a more challenging 'uncooperative DTI dataset' with increased number of artefactual DWIs and outlier slices. The lowest ROC AUC value in these simulations was 0.97 with SNR 8 showing that SOLID is capable of handling multiple artefactual DWIs per shell. E-F) visualizes a similar dataset with a higher *b*-value. The overall decrease in SNR due to the higher *b*-value resulted in a slightly lower ROC AUC value of 0.95 for the lowest SNR, and largest differences between the suggested $t_L =$ 3.5 and the optimal modified Z-score threshold. The lowest PRC AUC value was 0.84 indicating that the number of false positive findings in these simulations was overall low (Supplementary Fig. 1). A further investigation for -10% and -50% signal decreases (Supplementary Fig. 2) demonstrated that SOLID was well capable of identifying -50%deviations with the smallest ROC and PRC AUC values being 0.97 and 0.91, respectively whereas for the 10% decrease values were 0.91 and 0.55 indicating an increase in the number of false positive findings.

The subject motion between volumes might challenge the detection as SOLID performs a slicewise comparison across DWIs. Based on our investigations with relatively large 5-degree rotations (Fig. 3 and Supplementary Fig. 3), SOLID performed well even in the presence of these geometrical misalignments between DWIs with the lowest ROC AUC value being approximately 0.95. The lowest recorded PRC AUC was 0.80.

3.1.2. Comparison with existing algorithms for slicewise outlier detection

In Fig. 4A–D) we illustrate an outlier-scenario that can be problematic if the algorithms in DTIPrep and DTI Studio would be used to specify the outlier slices (currently they only flag the entire DWI volume if it contains slicewise outliers). For this example, only one DWI was artefactual. A coronal view of the DWI was used to visualize the positioning of adjacent and interleaved outlier slices (Fig. 4A). All tests were carried out using default settings for each algorithm.

In Fig. 4 B) we show that outlier detection based on the normalized cross-correlation of adjacent slices within DWIs found all outliers correctly but resulted in false positive findings in slices adjacent to true outliers. In C) we show SOLID results for the same example with all outliers correctly identified as full outliers with the modified Z-score being larger than the upper threshold ($\zeta_{k,l} > 10$) and false positives were found only in the inferior part of the volume, where only a few voxels contained signal from the brain. In D) we used the same example to visualize the problems of morphological closing-based detection where both adjacent outliers in the superior part of the brain and interleaved signal increase slices in the middle are not or incorrectly detected.

In Fig. 4 E) we compared the sensitivity and specificity of SOLID and FSL EDDY. SOLID is shown to be able to achieve at least similar specificity for the default setting of FSL EDDY (standard deviation threshold of 4), and to have an overall higher sensitivity to a given specificity, resulting in a 9% higher ROC and 2% higher PRC AUC values. The default threshold used in EDDY was well suited for this data as the improvement of the specificity and decrease of the sensitivity were modest with higher threshold values. The PRC AUC values were 0.87 and 0.85 for SOLID and EDDY respectively. The full PRC profiles for both algorithms are reported in Supplementary Fig. 4. The optimal threshold for EDDY was 2.4 σ or 40 σ based on ROC curve and PRC profiles respectively. The latter threshold was rather high and deviated substantially from the recommended setting (4 σ) where SOLID and EDDY are comparable.

3.1.3. SOLID-informed estimation: the effect of introducing uncertainty

Fig. 5 shows that the bias in tensor derived estimates (rows) FA, MD, KA, and the kurtosis tensor model residual can be significantly reduced, even in the presence of multiple outliers with large intensity deviations if outlier certainty can be estimated correctly and is used as a weight in model estimation. The y-axis of subfigures states the deviation from the GT CC signal and the x-axis the simulated certainty of the outlier signals, i.e. SOLID weights (eq. 5) from 1 to 0. This illustrates the effectiveness of the data certainty principle by remediating the estimation and decreasing the model residuals in the presence of outliers. The number of outliers per shell (columns) is varied from no outliers on the left to 16 outliers on the right. The color in each subfigure represents the median deviation from the corresponding GT measure or the residual value calculated over 1000 noise iterations.

Both signal decrease and increase outliers caused increasing deviation from the baseline (indicated with blue arrows) as the outlier magnitude increased. This simulation encompasses the whole gradual intensity band region from the middle with full outlier intensity to the edge regions, where, due to partial volume effects, only a fraction of the signal is from an outlier, and the remaining signal originates from 'normal' data. Decreasing the certainty of outlier data from 1 to 0 shows that signal deviations corresponding to different regions of the gradual intensity band can be handled by proper downweighting, and that small deviations in the signal do not require full data rejection to obtain a reliable model estimate. The deviation from the ground truth values increased for all estimates and model residuals with the decreasing SNR values and were higher in GM than in WM. The kurtosis tensor in particular was affected



Fig. 2. Sensitivity – specificity profiles (ROC curves) of SOLID for different 30-DWI simulations with four SNR values (8, 16, 32, Inf). The outlier slice signal was decreased by 100% (left) or increased by 50% (right). A-B) one outlier slice in one artefactual DWI with *b*-value 1000 s/mm². C-D) five outlier slices in eight artefactual DWIs with *b*-value 1000 s/mm². E-F) five outlier slices in eight artefactual DWIs with *b*-value 2000s/mm². Random selection of slice locations and artefactual DWIs was repeated 1000 times. The optimal thresholds are reported in each figure in the 'optimal Z-score' column.



Fig. 3. Sensitivity – specificity profiles (ROC curves) of SOLID for 30 DWI simulations with four SNR values (8, 16, 32, Inf) with *b*-value 1000 s/mm². The outlier signal was A-B) decreased by 100% or C-D) increased by 50%. A 5-degree rotation to artefactual DWI was induced around A,C) AP-axis and B,D) LR-axis. The random selection of slice locations and artefactual DWIs was repeated 1000 times. Examples of rotations and outlier locations are shown in each subfigure. The optimal thresholds are reported in each figure in the 'optimal Z-score' column.

by noise effects.

Lines indicated by the blue arrows also represent situations where SOLID would have detected false positive outliers i.e. the signal is not deviated from the ground truth but SOLID downweights it nevertheless. These lines show, that even with 16 outliers out of 60 DWIs per shell residual, FA, and MD estimates would not deviate from the baseline (the point with zero signal deviation on y-axis) and SOLID weight of 1 on xaxis. However, KA estimate would be visually affected at 8 outliers per shell and more notably with 16 outliers in the case of false positive findings.

In the extreme case of multiple outliers with large signal deviation, we observed that the design matrix (Batchelor et al., 2003) is slightly ill-conditioned thus noise propagates on each iteration step and the IWLLS estimator fails to converge to a finite solution. This is visualized in FA estimate 16 outlier case with -100% signal decrease and small SOLID downweighting in the subfigure left top corner. As the IWLLS estimator does not converge, the algorithm returns the initial LLS estimate which has been affected only once by the ill-conditioned design matrix.

Decreasing outlier certainty or intensity has similar effect and the IWLLS estimator converges to a finite solution.

It should be noted that in practice only large outlier deviations lead to a full SOLID downweighting factor of 0, thus areas shown in the top and bottom of the left side of each subfigure are not likely to be observed when applying SOLID. This means that SOLID can maintain the least squares problem mathematically well-conditioned even in the presence of 16 outliers per shell of 60 DWIs. For results in simulated GM signals with SNRs 16, 32, and 64, as well as WM signals with SNRs 16, 64, please see Supplementary Figs. 5–9.

Fig. 6 shows the influence of outlier signals in gradual intensity bands on different tensor derived estimates, and how SOLID-informed estimation can improve the results. Artefactual DWIs were rotated 5° around the left-right axis to simulate subject motion, resulting in gradual intensity bands after registration. Two adjacent slicewise outliers were introduced in both DWIs. Fig. 6 A) shows the location of the gradual intensity band on an axial DWI slice and Fig. 6 B) highlights the gradual intensity band in the FA image with the green line representing the edges of the FA >



Fig. 4. The outlier detection comparison of different tools. A) the same example of an extremely problematic case is shown twice to make the comparison of different tools in the first two rows easier. B) Detection based on normalized cross-correlation produces multiple false positives, indicated with the red arrows (all detections are colored in black). C) SOLID detects all outliers correctly and finds few false positives in one inferior slice of the volume because of the low number of tissue voxels. D) Detection based on morphological closing operation fails to detect multiple adjacent outliers with *i*) increased- and *ii*) decreased signal intensity, and *iii*) incorrectly positions interleaved signal increase outliers. E) 1000 samples of 30 DWIs with *b*-value 1000 s/mm² were used to compare the sensitivity and specificity of SOLID and FSL EDDY. Both the modified Z-score and standard deviation thresholds increase from right to the left. The color scale of the SOLID curve visualizes the linear decrease of the weight between the lower and the upper thresholds.

0.25 region that was used in the subsequent analysis. In Fig. 6 C) we compare SOLID weights to outlier locations identified with a voxelwise robust REKINDLE algorithm with the default kappa value of 6 to emphasize the problematic gradual intensity band edge, especially in the deep gray matter regions. Fig. 6 D) and E) depict the histograms of deviations from the GT for the FA and the first eigenvector (FE) and show that LLS and IWLLS estimators are severely affected by outliers, whereas IWLLS + SOLID and REKINDLE estimators give results closer to the GT. Fig. 6 F) and G) shows the spatial variation of the FA and FE deviations. While visual differences between the REKINDLE and the SOLID + IWLLS

results are modest, based on the outlier profiles at the gradual intensity band edge in C), REKINDLE would benefit from SOLID information as well. A detailed comparison of differences in FA estimates between EDDY (which uses signal prediction before estimation) and SOLID-based informed estimation (which incorporates missing data as uncertainty) is shown in Supplementary Fig. 10.

3.1.4. Real data experiments

To elucidate the data/certainty principle on in-vivo MRI data and further motivate the thresholding convention presented in eq. (4), Fig. 7



Fig. 5. Differences between simulated diffusion and kurtosis tensor measures FA, MD, KA, and model residual (rows of subfigures) as function of the number of outliers per shell 0, 1, 2, 4, 8, and 16 out of 60 (columns of subfigures). Pixels in each subfigure represent a median deviation from the ground truth measure or the residual within the corpus callosum voxel calculated over 1000 noise iterations with SNR 32. In the subfigures, the outlier intensity is decreasing from bottom to top on the y-axis and along the x-axis the outlier certainty decreases from left to right. The baseline with no signal outlier signal intensity (indicated with blue arrows) is shown as reference. The color scale of FA and MD is from zero to maximum deviation, KA from zero to 95% percentile deviation as the maximum deviation would have saturated subfigures, and residual from minimum to maximum.

shows instances of an adult subject without visible intensity artefacts on the first row, and a neonatal subject with severe subject motion related artefacts on the second row. As both subjects were acquired with axial slices, a sagittal view of DWIs before and after subject motion correction (Fig. 7A) indicates the spatial locations of these artefacts on the neonatal data. If a standard IWLLS estimator would be used on these DWIs to compute the diffusion tensor, the weights of the final iteration cycle would be analogous to the diffusion weighted signal simulated by the estimated tensor at the end of the previous cycle. Investigation of the weights throughout the iteration process on neonatal data shows that the final model estimation is highly affected by outliers (Fig. 7B) as the weights of the outlier slices deviate significantly from zero. By applying SOLID weights calculated with $t_L = 0$ in the IWLLS estimator (eq. 5), the neonatal iteration cycle is improved as outliers are correctly downweighted, but the adult data is influenced by incorrect weights because the lower threshold t_{l} is too low to account for the normal variation of the slicewise intensity (Fig. 7C). By increasing the lower threshold to 3.5 this is avoided, and both the final weights of the adult and the neonatal subject are visually more plausible (Fig. 7D).

Fig. 7 E) shows the evaluation of SOLID on 54 neonatal subjects which were visually inspected for slicewise outliers. While the original ROC AUC value of 0.92 indicates a good agreement between automated and visual slicewise outlier detections, the original PRC AUC of 0.47 implied that there could be many false positive findings (Fig. 7 F). The initial SOLID results were revised to visually confirm if the finding was really a true or a false positive. The revised AUC values were 0.97 and 0.79 for ROC and PRC, respectively, indicating the fact that technician had missed multiple, more subtle outliers during the original visual inspection of the data. When inspecting the detection results, we noted that the most drastic outliers were detected both visually and using SOLID, but SOLID was also able to find artefactual slices that were not recognized during visual inspection. Fig. 7 G) shows an example of the same axial slice and

the same gradient direction from two subsequent acquisitions of the same subject: the upper slice showed no anomalies and had a modified Z-score of 0.4, whereas the lower slice had local intensity decreases (indicated with red arrows) and a modified Z-score of 8.9. These anomalies arose due to suboptimal hardware behavior and were not detected in adjacent slices, nor DWIs acquired in any other direction or in non-diffusion weighted images, thus excluding anatomical lesions. Such local abnormalities could be easily missed in visual inspection since observing them visually from sagittal and coronal views is challenging.

4. Discussion

Subject motion or hardware issues during the acquisition of diffusionweighted MRI data can greatly challenge the unambiguous characterization of the underlying physical diffusion process. Handling such errors manually is challenging and time consuming. Existing tools (Liu et al., 2010; Li et al., 2013; Andersson and Sotiropoulos, 2016; Jiang et al., 2009) that aim to identify and handle slicewise outliers showed room for improvement in accuracy, ease of use, and speed (e.g. Fig. 4). In this work, we developed i) an approach called SOLID to robustly detect both signal decrease and increase artifacts from raw DWIs and ii) a framework for the use of this information to estimate the certainty of measurements and robustly guide tensor estimation algorithms.

The main purpose of the SOLID-based estimation framework is not to recover data – as one can argue that essentially not more information is added. If a measurement is corrupted, information along that gradient direction is lost resulting in a 'missing or uncertain data problem'. While this missing point can be simulated from the remaining measurements (Andersson et al., 2016; Morris et al., 2011) it does not increase the information available. We used both comprehensive simulations as well as a large amount of real human data to test and validate the performance of SOLID, showing a good performance on typical diffusion MRI datasets,

V. Sairanen et al.



NeuroImage 181 (2018) 331-346

Fig. 6. A practical example of the gradual intensity band artefact affecting diffusion tensor estimators. A) Location of the gradual intensity band on an axial DWI slice. B) The FA values of the same slice with the green line showing the edges of the FA>0.25 region, and the blue rectangle highlighting the gradual intensity band. C) The SOLID weights $S_l(\mathbf{x}')$ of the artefactual slice are shown on the colormap along with the FA>0.25 edges. REKINDLE outliers, shown with dark green '+' signs, emphasize the challenging regions for voxelwise algorithms near the gradual intensity band edge. D-E) Histograms of FA and FE deviations from the ground truth within the FA>0.25 region for LLS, IWLLS, SOLID informed IWLLS and REKINDLE algorithms. F-G) Spatial maps for FA and FE deviations within the gradual intensity band indicated by the blue square in B).

particularly on clinical datasets with a low number of acquired DWIs.

4.1. Outlier detection

Outlier detection in SOLID is not dependent on any model estimation or prediction and there is no hard requirement on the amount of data gathered or on the specific directions of used gradient vectors. The only practical restriction is that if over half of DWIs have a corrupted slice at exactly the same spatial location, the algorithm cannot identify the outliers due to the median based detection. Although SOLID assumes that observations used to determine the modified Z-score (eq. 2) follow a normal distribution, our simulations (Figs. 2, Fig. 3, and Fig. 4) and real data examples (Fig. 7) show that potential deviations from this assumption do not hinder the outlier detection drastically. Similar assumptions are made within previously published tools as well (Liu et al., 2010; Li et al., 2013; Andersson and Sotiropoulos, 2016), and we argue that in general the differences in the intensity metric investigated between outliers and normal slices in diffusion MRI data are so apparent



Fig. 7. Effect of applying SOLID weights on the in-vivo MRI data and validation of the outlier detection algorithm. A1) Non-corrected and A2) motion corrected DWI of the adult shows no visual intensity errors whereas the motion corrected neonatal data has multiple outlier slices slightly interpolated into adjacent slices. B) Weights of the final iteration cycle of the standard IWLLS estimator are undesirably high in outlier slices for the neonate. C) SOLID weights calculated with lower threshold $t_L = 0$ causes unnecessary downweighting of good quality data (adult) whereas the weights of the neonatal data are more appropriate. D) The lower threshold was increased to 3.5 to reduce downweighting considering the normal variation in the data, yielding proper weights on the final iteration cycle for both subjects. E) The sensitivity – specificity profile and F) the precision – recall curve of SOLID. Original curves were calculated based on the initial visual detection of outlier slices in the neonatal dataset whereas Revised curves were based on the second visual inspection of SOLID results to pinpoint possible new true positive findings that were not identified during the initial labelling. G) An example of a local outlier case which was missed by the technician but not by SOLID. The upper slice shows no signal deviations thus excluding pathologies whereas the lower slice which had the same gradient orientation from the same subject with approximately the same slice location shows two local signal decreases indicated with red arrows.

that the influence of the violations of this assumption remains limited.

We investigated a possible shortcoming of SOLID related to the initial geometrical misalignment of images due to severe subject motion (Fig. 3A and B). If the subject moves their head during the acquisition between DWIs, this could lead to a situation where a slice containing air cavities near the ears is compared with a slice in another DWI that did not, or did only partially, cover such air-tissue interfaces. While such misalignments slightly decreased the performance of SOLID (Figs. 2 and 3), the impact is likely limited by the informed estimation framework. Both the upper and the lower extremes of the brain were the most vulnerable to these rotation artefacts as they lead into situations where a slice did not have any signal from brain tissue in one DWI whereas other volumes had. A possible solution to this is to set a threshold on the minimal number of tissue voxels per slice, as done in FSL EDDY. However, when we compared SOLID results to FSL EDDY on DWI simulations with subject motion (Fig. 4 E) we found that while EDDY and SOLID have

similar maximal specificity, SOLID had overall better sensitivity resulting in 9% higher ROC and 2% higher PRC AUC values. The lower sensitivity of EDDY is likely caused by the hardcoded setting for the minimal number of tissue voxels per slice, which is not employed for SOLID. We therefore opted to avoid a voxel threshold to increase sensitivity at the risk of slightly lower specificity. EDDY recognized multiple false positives as well which was likely due to the large number of outliers simulated at the same slice position. It should be noted that with a significantly higher standard deviation threshold than the default value, EDDY resulted in a better positive prediction value than SOLID, but the overall PRC AUC value was lower.

Since we did not introduce the requirement of a minimum number of voxels needed for the analysis, we found a notable discontinuity in the ROC curves of our outlier detection simulation tests (i.e. a 'drop' of approximately 2.5% in sensitivity, Figs. 2 and 3) even with infinite SNR near 50% specificity. This was a result of the random positioning of

outlier slices: in some cases, the slice was in the lower or in the upper extreme of the DWI volume which had only a few or no voxels with brain tissue. Variance of a small number of voxels could fall to the range of noise variance thus such slices were detectable only with extremely low modified Z-score thresholds. Signal increase artefacts, however, resulted in a detectable change in variance.

While we only presented results with variance as a slicewise metric in the SOLID framework in this paper, we did do preliminary tests with the mean, median, and MAD slicewise intensity as well, and benchmarked them against visually detected outliers in the neonatal dataset (Fig. 8). These metrics resulted in slightly different modified Z-scores (traditional Z-scores for median and MAD as they induced computational challenges with the modified Z-score) for outlier slices. Based on the ROC AUC values and curves, variance and mean produced nearly identical results whereas MAD and median metrics were clearly inferior to them. We also studied more complex metrics that take more information of the slicewise histograms into account, i.e. with Kolmogorov-Smirnov statistics and Mutual Information of slices as inputs to SOLID. Since they did not significantly improve the detection (results not shown) but increased both the complexity of SOLID and computational time, and simpler metrics such as variance had already shown good agreement with visual detections, we did not explore them further.

Although we suggest to use a lower threshold of $\xi'_{l}(\mathbf{x}') = 3.5$ for variance based SOLID (motivated by our simulation results and previous literature (Iglewicz and Hoaglin, 1993)) this is an arbitrary threshold and could be adjusted if necessary. For example, in the case of noisy data with multiple outliers the optimal lower thresholds were $t_L \approx 1.9$ for the data of 54 neonatal subjects (Fig. 7 E), $t_L \approx 2.1$ for higher *b*-value data (Fig. 2E and F), and $t_L \approx 2.5$ for subject motion simulations (Fig. 3). However, changes in the lower threshold might not affect the final model estimate significantly due to the linear downweighting approach (eq. 4).

4.2. Comparison with existing algorithms for slicewise outlier detection

From the multiple existing tools (Zhou et al., 2011; Lauzon et al., 2013; Oguz et al., 2014; Jiang et al., 2006; Pannek et al., 2015), we selected three algorithms the most similar to SOLID for closer inspection and testing: DTIPrep (Liu et al., 2010), DTI Studio (Li et al., 2013) – of which the inter-slice intensity discontinuity (ISID)-metric has been

applied in the methods of correcting for out-of-plane inconsistency in the acquired images (Marami et al., 2016; Marami et al., 2016), and FSL EDDY (Andersson and Sotiropoulos, 2016; Andersson et al., 2017). As we showed in Fig. 4, the fast-to-compute slicewise variance metric employed for outlier detection in this work was capable of competing with these more complex algorithms. The algorithms in DTIPrep and DTI Studio searched for outliers within one DWI. While this avoids the problem of subject motion between DWIs, it could lead to multiple false positive findings (Fig. 4). In the case of DTIPrep and DTI Studio, false positive findings could lead to the rejection of the whole DWI volume. For example, (Marami et al. 2016), reported that they excluded the whole DWI if more than 15% slices were outliers based on the ISID-metric. This means that a small number of outliers could result in the rejection of non-corrupted slices, which could in turn affect model estimates or approaches that aim to perform slice-level motion correction. By incorporating the SOLID downweighting scheme, unnecessary data loss could be avoided, and the remaining good measurements could still be exploited.

On the other hand, the FSL EDDY outlier detection is more similar to SOLID in that it employs the information across DWIs. However, it is based on the more complex prediction of the signal and sets restrictions on how data should be acquired, i.e. the minimum allowed number of gradient directions, the distribution of gradients, and the localization of measurements on exact q-space shells. For example, based on the user guide of EDDY, using it on our neonatal dataset acquired with 15 direction over-plus gradient scheme would not be advised. SOLID lifts these restrictions on the gradient distribution as it works with fewer directions, and is likely robust to small deviations in b-values (e.g. due to gradient nonlinearities) as the deviation from a signal with the predefined set bvalue might not be prominent. EDDY outlier detection currently sets a hard threshold on the number of voxels per slice that must contain a signal from tissue that could not be further decreased than 250 in the version tested, which is likely the reason for the overall lower sensitivity compared to SOLID in (Fig. 4). It should be noted that FSL EDDY provides, apart from the outlier detection, a complete framework for correcting geometrical artefacts (Andersson and Sotiropoulos, 2016; Andersson et al. 2003, 2017; Jenkinson et al., 2012) and SOLID could be used together with the geometrical distortion correction framework (or any other motion/distortion correction strategy) or even serve as input to the EDDY outlier detection.



Fig. 8. The initial benchmarking of different SOLID metrics: variance, mean, MAD, and median. A) Sensitivity – specificity and B) precision-recall profiles were calculated from visually detected outliers from the neonatal dataset.

Finally, we did not opt to replace outliers with predictions but instead chose to incorporate it as data certainty in the model fitting. While the replacement of outlier data can avoid the occurrence of the gradual intensity band after processing, this means that after misalignment correction part of the signals in gradual intensity band voxels are from true measurements and part are predicted. It remains to be investigated how this combination of measurement- and prediction-errors will propagate into the tensor model estimations.

4.3. SOLID-informed estimation: the effect of introducing uncertainty

We selected typical GT signals from deep GM and CC to study how partial voluming in gradual intensity band would affect diffusion and kurtosis tensor derived measures, and how this could be corrected for by including the data uncertainty derived from SOLID (Fig. 5 and Supplementary Figs. 5–9). For both tissue types, the trends were similar: like expected, an increase in the intensity deviation or the number of outlier signals per shell lead to a larger deviation from the GT in all tensor derived measures and model residuals. Concomitantly, both the signal decrease and increase artefacts could be ameliorated by introducing data uncertainty as a weight during IWLLS estimation (eq. 5).

In a few extreme cases of multiple outliers per shell, depending which random gradient directions were picked, the IWLLS estimator failed to converge into a finite solution. This numerical problem becomes visible in the upper left corner of the FA subfigure of Fig. 5 with 16 outliers per shell: the FA deviation for outliers with 100% signal decrease is smaller than for outliers with a decrease of 90%. We found that, if the used IWLLS estimator did not converge within a predefined number of iterations, it would return the initial LLS estimate instead. The detailed inspection of the IWLLS iterations revealed that the condition number of the iteratively reweighted design matrix started to rapidly increase, causing cumulative error propagation during each iteration step (Sairanen et al., 2017; Batchelor et al., 2003; Skare et al., 2000).

By adapting the data uncertainty (SOLID weights) in this simulation, we gradually decreased the weight of outlier signals. This kept the design matrix well-conditioned, avoided the numerical problems, and significantly reduced the bias in estimates. For the sake of completeness, even some unrealistic situations were investigated where the data certainty was high while the intensity deviation was large: Fig. 5 the left most corners in each subfigure. In practice, such slices with -100% or +50% deviations are generally easily detected by SOLID and their weights are adjusted accordingly. This also means that the zero-deviation row i.e. 'no outliers' situation would practically never be fully downweighted. On the other hand, the zero row also shows that SOLID informed model estimation is quite robust for false positive findings as well.

Additionally, we did a comparison between LLS, WLLS, voxelwise robust REKINDLE (Tax et al., 2015), and SOLID-informed IWLLS diffusion tensor algorithms (Fig. 6). Both the robust and informed methods produced similar improvements to the non-robust algorithms. When inspecting the gradual intensity band more closely, it becomes clear that voxelwise methods fail to detect some outlier voxels. In the case of REKINDLE, a simple adjustment of the κ -parameter could likely correct this, but adjusting parameters for each subject separately would be time consuming. The optimal solution would be to inform robust estimators like REKINDLE with the SOLID weights; any voxels with $\xi'_{l}(\mathbf{x}') > t_{U}$ could be flagged as outliers in the initial fit to ease and speed up the robust estimation process.

4.4. Real data experiments

We evaluated SOLID on adult dMRI data which did not visually suffer from any intensity artefacts related to subject motion, and on 54 dMRI datasets of neonatal subjects which had low SNR and multiple visible motion artefacts. The purpose of using the neonatal dataset was to obtain information how SOLID would perform in the presence of variety of motion artefacts, as it is challenging to simulate this in a realistic manner. Our simulations were limited to such subject motion or hardware issues that cause a full slice wise artefact. In real acquisitions, however, motion is likely to happen continuously (Jiang et al., 2009; Fogtmann et al., 2014; Marami et al., 2016) and could corrupt the signal between slices or result in different rotations for each slice within the DWI volume.

One neonate and one adult were highlighted to visualize and motivate the lower modified Z-score threshold selection t_L (eq. (4)) in Fig. 7 A–D. The adult data was from a recent study and consisted of relatively modern acquisition scheme with 64 gradient directions and high SNR whereas the unique neonatal data was from a more dated study with inferior gradient scheme and lower SNR. The variability of the datasets tested along with the variety of simulations shows the versatility of SOLID. While SOLID missed some of the visually labeled outliers (still resulting in an adequate ROC AUC value of 0.92 in Fig. 7 E), it detect multiple milder outliers that were ignored by the technician. The initial visual inspection revealed 355 outliers out of all 102960 slices (0.3%) whereas the revised visual inspection based on SOLID results revealed 953 outliers (0.9%) i.e. the technician missed 598 outlier slices. The number of false negative findings due to outliers being in the upper or the lower extreme of the DWI volume was 98.

SOLID output was visually inspected to confirm true and false positive findings to generate revised ROC and PRC curves. As SOLID was able to find multiple outliers that were missed in the initial inspection ROC AUC increased from 0.92 to 0.97 and PRC AUC 0.47 to 0.79. For example, Fig. 7 G the upper slice shows no visible artefacts whereas the lower slice acquisition endured hardware issues leading to local signal decreases (indicated with red arrows). These local deviations were enough to alter the slicewise variance metric which SOLID could detect.

4.5. Current implementation of SOLID

SOLID as well as the SOLID-based informed estimation framework will be distributed with the Matlab toolbox ExploreDTI (Leemans et al., 2009) and an interactive tool will be made available (https://github. com/vilsaira/SOLID) to provide a fast overview of suspicious slices within the data, and as such will greatly facilitate quality assessment in large group studies. To facilitate SOLID-based estimation in other software than ExploreDTI, non-transformed and transformed modified Z-score volumes ($\xi(\mathbf{x})$ and $\xi'(\mathbf{x'})$) and final SOLID-weight volumes $S_l(\mathbf{x'})$ will be given as outputs. The non-transformed modified Z-scores can be used to investigate geometric transformations obtained from other registration software. Subsequently, the transformed weights can be input to estimation software e.g. Maastricht Diffusion Toolbox (Harms et al., 2017).

4.6. Future aspects of SOLID

The sensitivity and specificity of SOLID could potentially be further improved by exploiting additional within-DWI and across-DWI information. Adding within-DWI information might improve the ROC AUC and PRC values in the presence of large rotational subject motion, but further work is needed to investigate how this additional information is optimally included while maintaining high performance in the challenging cases illustrated in Fig. 4. For example, the size of the 'slice neighborhood' and the coupling of the within-DWI and the across-DWI should be further explored when including this metric. Additional across-DWI information can be included by exploiting the angular relation between DWIs, i.e. by angular matching (St-Jean et al., 2016). When comparing the slicewise metric, more emphasis can be put on DWIs that are closer in terms of their angular neighborhood. This could be beneficial when performing outlier detection on DWIs acquired with extremely high b-values with only a little signal at locations where the fiber population is approximately perpendicular to the encoding direction. While this would reintroduce restrictions on the minimum angular

sampling, it is generally recommended to acquire very high b-value shells with dense angular sampling. Including the additional within-slice- and across-slice neighborhood information could potentially extend the usability of SOLID to such research setups that are beyond typical clinical research settings. Preliminary qualitative results on high b-value (6000 s/mm²) data suggest that SOLID as presented in this work can pick up outlier slices without extensively reporting false positives.

The SOLID framework could also be extended to other models and representation approaches for dMRI (Tournier et al. 2004, 2007; Tuch, 2004; Wedeen et al., 2005) and to other estimators beyond least squares (Veraart et al., 2011; Andersson, 2008; Liu et al., 2016). The gradual downweighting approach could also be applied based on results from other tools that consider more complex within-volume motion (Andersson et al., 2017; Fogtmann et al., 2014; Ferrante and Paragios, 2017; Jiang et al., 2009; Marami et al., 2016; Marami et al. 2016, 2016). Similar to our implementation with the modified Z-score, FSL's EDDY produces standard deviation scores that could be used as estimates for data reliability and traced through image registration and resampling steps to provide weights during the model estimation.

Currently, SOLID is designed to detect slice wise outliers of any kind (including but not limited to those arising from subject motion). While we simulated artefacts as complete slicewise artefacts that could for example have been caused by motion, motion in real data can be more complex. The benchmarking with 54 neonatal subjects complements the simulation results in that it exposes SOLID to a more complex set of motion artefacts, but it is more challenging to establish a ground truth. Future work can be addressed to evaluating outlier detection methods in more realistic simulations.

5. Conclusion

We presented a fast and automatic tool coined SOLID to detect and handle slicewise outliers (both the signal decrease and increase) in multishell diffusion MRI data. SOLID relaxes the requirements on the amount and angular resolution of data needed for successful outlier detection compared to other tools. Adding SOLID to existing preprocessing pipelines is straightforward, and while the standard settings worked well for the scenarios investigated, the user can further tune the intuitive parameters if desired.

SOLID was evaluated on the extensive sets of simulated outlier cases having different clinical b-values and gradient schemes, and further validated on 156 sets of 15 DWIs from 54 different neonatal subjects which were visually inspected for slicewise outliers. The performance of SOLID in simulations was higher or at least as good compared to three previously published slicewise outlier detection algorithms that were generally more complex and computationally demanding. The challenging evaluation in in-vivo neonate data emphasized that SOLID was in good agreement with the visual inspections, robust towards subject motion related geometrical misalignments, and able to identify milder outliers missed by the technician showing its potential to quickly highlight problematic volumes and slices in large population studies.

The presented framework utilized the SOLID information during diffusion and kurtosis tensor model estimation by including it as 'uncertainty' of the measured data. The framework was able to ameliorate estimation in the presence of outliers, and to keep the estimation process numerically stable in extreme cases with a large number of outliers, by restraining the rapid increase of the condition number in iteratively reweighted estimation.

Acknowledgements

The authors thank Derek Jones for the valuable discussions and comments on the manuscript and Antti Kuronen for supporting the usage of the Alcyone computing cluster of the University of Helsinki Department of Physics for the computational tasks in this study. V.S is supported by Helsinki University Hospital, Juselius Foundation, Finnish Cultural Foundation, the Academy of Finland grant (288220 and 276523). C.T. is supported by a Rubicon grant (680-50-1527) from the Netherlands Organisation for Scientific Research and Wellcome Trust grant (096646/Z/11/Z). A.L. is supported by VIDI Grant 639.072.411 from the Netherlands Organisation for Scientific Research.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.07.003.

References

- Andersson, J.L., Skare, S., 2010. Image distortion and its correction in diffusion MRI. In: Diffusion MRI: Theory, Methods, and Applications, pp. 285–302.
- Andersson, J.L.R., et al., 2016. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. Neuroimage 141, 556–572.
- Andersson, J.L.R., 2008. Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: why, how and but. Neuroimage 42 (4), 1340–1356.
- Andersson, J.L.R., et al., 2017. Towards a comprehensive framework for movement and distortion correction of diffusion MR images: within volume movement. Neuroimage 152, 450–466.
- Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage 20 (2), 870–888.
- Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for offresonance effects and subject movement in diffusion MR imaging. Neuroimage 125, 1063–1078.
- Andersson, J.L.R., Sotiropoulos, S.N., 2015. Non-parametric representation and prediction of single- and multi-shell diffusion-weighted MRI data using Gaussian processes. Neuroimage 122, 166–176.
- Basser, P.J., Mattiello, J., Lebihan, D., 1994. Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson., Ser. B 103 (3), 247–254.
- Batchelor, P.G., et al., 2003. Anisotropic noise propagation in diffusion tensor MRI sampling schemes. Magn. Reson. Med. 49 (6), 1143–1151.
- Baum, S.H., Stevenson, R.A., 2016. Commentary: visual fixation in human newborns correlates with extensive white matter networks and predicts long-term neurocognitive development. Front. Neurosci. 4824–4829, 10(MAY).
- Le Bihan, D., et al., 2006. Artifacts and pitfalls in diffusion MRI. J. Magn. Reson. Imag. 24 (3), 478–488.
- Chang, L.C., Jones, D.K., Pierpaoli, C., 2005. RESTORE: robust estimation of tensors by outlier rejection. Magn. Reson. Med. 53 (5), 1088–1095.
- Chang, L.C., Walker, L., Pierpaoli, C., 2012. Informed RESTORE: a method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. Magn. Reson. Med. 68 (5), 1654–1663.
- Collier, Q., et al., 2015. Iterative reweighted linear least squares for accurate, fast, and robust estimation of diffusion magnetic resonance parameters. Magn. Reson. Med. 73 (6), 2174–2184.
- Cook, P., et al., 2006. Camino: open-source diffusion-MRI reconstruction and processing. In: 14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine, vol. 14, p. 2759.
- Ferrante, E., Paragios, N., 2017. Slice-to-volume medical image registration: a survey. Med. Image Anal. 39, 101–123.
- Fogtmann, M., et al., 2014. A unified approach to diffusion direction sensitive slice registration and 3-d DTI reconstruction from moving fetal brain anatomy. IEEE Trans. Med. Imag. 33 (2), 272–289.
- Froeling, M., et al., 2017. ?MASSIVE? brain dataset: multiple acquisitions for standardization of structural imaging validation and evaluation. Magn. Reson. Med. 77 (5), 1797–1809.
- Hämäläinen, S., et al., 2017. Bilingualism modulates the white matter structure of language-related pathways. Neuroimage 152, 249–257.
- Harms, R.L., et al., 2017. Robust and fast nonlinear optimization of diffusion MRI microstructure models. Neuroimage 155, 82–96.
- Heemskerk, A.M., et al., 2013. Acquisition guidelines and quality assessment tools for analyzing neonatal diffusion tensor MRI data. Am. J. Neuroradiol. 34 (8), 1496–1505.
- Iglewicz, B., Hoaglin, D.C., 1993. How to Detect and Handle Outliers, Jenkinson, M. et al., 2012. Fsl. NeuroImage 62 (2), 782–790.
- Jezzard, P., Barnett, A.S., Pierpaoli, C., 1998. Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. Magn. Reson. Med. 39 (5), 801–812.
- Jenkinson, Mark, et al., 2012. Fsl. Neuroimage 62.2, 782-790.
- Jiang, H., et al., 2006. DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking. Comput. Meth. Progr. Biomed. 81 (2), 106–116.
- Jiang, S., et al., 2009. Diffusion tensor imaging (DTI) of the brain in moving subjects: application to in-utero fetal and ex-utero studies. Magn. Reson. Med. 62 (3), 645–655.
- Jones, D.K., 2008. Studying connections in the living human brain with diffusion MRI. Cortex 44 (8), 936–952.
- Jones, D.K., Cercignani, M., 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR Biomed. 23 (7), 803–820.

V. Sairanen et al.

Jones, D.K., Horsfield, M.A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn. Reson. Med. 42 (3), 515–525.

- Kennis, M., et al., 2016. Choosing the polarity of the phase-encoding direction in diffusion MRI: does it matter for group analysis? Neuroimage: Clinica 11, 539–547.
- Knutsson, H., Westin, C., 1993. Normalized and differential convolution: methods for interpolation and filtering of incomplete and uncertain data. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June(16–19), pp. 515–523.
- Langen, M., et al., 2012. Fronto-striatal circuitry and inhibitory control in autism: findings from diffusion tensor imaging tractography. Cortex 48 (2), 183–193.
- Lauzon, C.B., et al., 2013. Simultaneous analysis and quality assurance for diffusion tensor imaging. PLoS One 8 (4).
- Leemans, A., et al., 2009. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine 17 (2), 3537.
- Li, Y., et al., 2013. Image corruption detection in diffusion tensor imaging for postprocessing and real-time monitoring K. Herholz. PLoS One 8 (10), e49764.
- Liu, J., Gasbarra, D., Railavo, J., 2016. Fast estimation of diffusion tensors under Rician noise by the EM algorithm. J. Neurosci. Meth. 257, 147–158.
- Liu, Z., et al., 2010. Quality control of diffusion weighted images. In: Proceedings of SPIE-the International Society for Optical Engineering. NIH Public Access, p. 76280J.
- Lucas, B.C., et al., 2010. The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software. Neuroinformatics 8 (1), 5–17.
- Mangin, J.F., et al., 2002. Distortion correction and robust tensor estimation for MR diffusion imaging. Med. Image Anal. 6, 191–198.
- Mangin, J.F., et al., 2001. Eddy-current distortion correction and Robust tensor estimation for MR diffusion imaging. Lect. Notes Comput. Sci. 2208, 186–194.
- Marami, B., Scherrer, B., Afacan, O., Erem, B., et al., 2016a. Motion-robust diffusionweighted brain MRI reconstruction through slice-level registration-based motion tracking. IEEE Trans. Med. Imag. 35 (10), 2258–2269.
- Marami, B., Scherrer, B., Afacan, O., Warfield, S.K., et al., 2016b. Motion-robust reconstruction based on simultaneous multi-slice registration for diffusion-weighted MRI of moving subjects. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Cham, pp. 544–552.
- Meer, P., et al., 1991. Robust regression methods for computer vision: a review. Int. J. Comput. Vis. 6 (1), 59–70.
- Morris, D., et al., 2011. Preterm neonatal diffusion processing using detection and replacement of outliers prior to resampling. Magn. Reson. Med. 66 (1), 92–101.
- Nielsen, J.F., Ghugre, N.R., Panigrahy, A., 2004. Affine and polynomial mutual information coregistration for artifact elimination in diffusion tensor imaging of newborns. Magn. Reson. Imag. 22 (9), 1319–1323.
- Norman, G.R., Streiner, D.L., 2007. Biostatistics: the Bare Essentials, PMPH-USA.
- Odish, O.F.F., et al., 2015. Microstructural brain abnormalities in Huntington's disease: a two-year follow-up. Hum. Brain Mapp. 36 (6), 2061–2074.
- Oguz, I., et al., 2014. DTIPrep: quality control of diffusion-weighted images. Front. Neuroinf. 8, 4.
- Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recogn. 29 (1), 51–59.
- Pannek, K., et al., 2015. Combined slicewise and voxelwise outlier detection in diffusion MRI data of preterm born infants. Ohbm.

- Pannek, K., et al., 2012. HOMOR: higher order model outlier rejection for high b-value MR diffusion data. Neuroimage 63 (2), 835–842.
- Pierpaoli, C., 2010. Artifacts in diffusion MRI. In Diffusion MRI: Theory, methods, and applications 303–317.
- Pierpaoli, C., Walker, L., 2010. TORTOISE: an integrated software package for processing of diffusion MRI data. Processing of Diffusion ... 51 (1), 2010.
- Poot, D.H.J., et al., 2010. Optimal experimental design for diffusion kurtosis imaging. IEEE Trans. Med. Imag. 29 (3), 819–829.
- Rohde, G.K., et al., 2004. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. Magn. Reson. Med. 51 (1), 103–114.
- Sairanen, V., et al., 2017. A novel measure of reliability in Diffusion Tensor Imaging after data rejections due to subject motion. Neuroimage 147, 57–65.
- Scelfo, D., et al., 2012. Automated detection, evaluation and removal of DWI-related artifacts. In Proc. Intl. Soc. Mag. Reson. Med 3553.
- Skare, S., et al., 2000. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. J. Magn. Reson. 147 (2) pp.340–52.
- St-Jean, S., Coupé, P., Descoteaux, M., 2016. Non Local Spatial and Angular Matching: enabling higher spatial resolution diffusion MRI datasets through adaptive denoising. Med. Image Anal. 32, 115–130.
- Stejskal, E.O., Tanner, J.E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J. Chem. Phys. 42 (1), 288–292.
- Stjerna, S., et al., 2015. Visual fixation in human newborns correlates with extensive white matter networks and predicts long-term neurocognitive development. J. Neurosci. 35 (12), 4824–4829.
- Tax, C.M.W., et al., 2017. Quantifying the brain's sheet structure with normalized convolution. Med. Image Anal. 39, 162–177.
- Tax, C.M.W., et al., 2015. REKINDLE: Robust extraction of kurtosis INDices with linear estimation. Magn. Reson. Med. vol. 73 (2), 794–808.
- Tax, C.M.W., Vos, S.B., Leemans, A., 2016. Checking and correcting DTI data. In: Diffusion Tensor Imaging. Springer, New York, NY, pp. 127–150. New York.
- Tournier, J.-D., Mori, S., Leemans, A., 2011. Diffusion tensor imaging and beyond. Magn. Reson. Med. 65 (6), 1532–1556.
- Tournier, J.D., et al., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage 23 (3), 1176–1185.
- Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage 35 (4), 1459–1472.
- Tuch, D.S., 2004. Q-ball imaging. Magn. Reson. Med. 52 (6), 1358–1372. Veraart, J., et al., 2013. Weighted linear least squares estimation of diffusion MRI
- parameters: strengths, limitations, and pitfalls. Neuroimage 81, 335–346.
- Veraart, J., Van Hecke, W., Sijbers, J., 2011. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. Magn. Reson. Med. 66 (3), 678–686.
- Vos, S.B., et al., 2017. The importance of correcting for signal drift in diffusion MRI. Magn. Reson. Med. 77 (1), 285–299.
- Wedeen, V.J., et al., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54 (6), 1377–1386.
- Zhou, Z., et al., 2011. Automated artifact detection and removal for improved tensor estimation in motion-corrupted DTI data sets using the combination of local binary patterns and 2D partial least squares. Magn. Reson. Imag. 29 (2), 230–242.