Published in final edited form as:

J Magn Reson Imaging. 2011 May ; 33(5): 1194–1208. doi:10.1002/jmri.22535.

Quality assessment of High Angular Resolution Diffusion Imaging data using bootstrap on Q-ball reconstruction

J. Cohen-Adad, PhD^{1,2}, M. Descoteaux, PhD⁴, and L.L. Wald, PhD^{1,2,3}

¹Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

²Harvard Medical School, Boston, MA, USA

³Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA, USA

⁴MOIVRE Centre, Department of Computer Science, Université de Sherbrooke, QC, Canada

Abstract

Purpose—To develop a bootstrap method to assess the quality of High Angular Resolution Diffusion Imaging (HARDI) data using Q-Ball imaging (QBI) reconstruction.

Materials and Methods—HARDI data were re-shuffled using regular bootstrap with jackknife sampling. For each bootstrap dataset, the diffusion orientation distribution function (ODF) was estimated voxel-wise using QBI reconstruction based on spherical harmonics functions. The reproducibility of the ODF was assessed using the Jensen-Shannon divergence (JSD) and the angular confidence interval was derived for the first and the second ODF maxima. The sensitivity of the bootstrap method was evaluated on a human subject by adding synthetic noise to the data, by acquiring a map of image signal-to-noise ratio (SNR) and by varying the echo time and the b-value.

Results—The JSD was directly linked to the image SNR. The impact of echo times and b-values was reflected by both the JSD and the angular confidence interval, proving the usefulness of the bootstrap method to evaluate specific features of HARDI data.

Conclusion—The bootstrap method can effectively assess the quality of HARDI data and can be used to evaluate new hardware and pulse sequences, perform multi-fiber probabilistic tractography, and provide reliability metrics to support clinical studies.

Keywords

HARDI; Q-Ball imaging; bootstrap; reproducibility; SNR; orientation distribution function

INTRODUCTION

Magnetic resonance diffusion-weighted imaging enables the mapping of axonal architecture in the central nervous system by exploiting the anisotropic properties of water diffusion in the white matter (1). Recently, High Angular Resolution Diffusion Imaging (HARDI) has been suggested as a means to overcome some limitations imposed by diffusion tensor imaging (DTI), notably in presence of crossing fibers (2). HARDI-based reconstruction methods such as Q-Ball imaging (QBI) have proven efficient in the detection of subtle axonal connections in the cerebrum (2) and spinal cord (3).

Whilst diffusion measurements convey a variety of useful information, they are also hampered by a very low signal-to-noise ratio (SNR). This especially holds true when going to HARDI techniques, where larger b-values are needed to depict low angle crossing fibers (4–6). The low sensitivity of diffusion measurements therefore motivates a more thorough evaluation of their reproducibility. In particular, having a reliability assessment for the orientation distribution function (ODF, the probability of water molecules to diffuse as a function of orientation) and its maxima is of major interest. In addition to aiding the optimization of acquisition methods, a measure of statistical confidence on the reconstructed ODF is useful for supporting clinical findings and for performing probabilistic tractography.

Multiple DTI-based methods have been proposed to assess the reproducibility of diffusion measurements (7–12). Most of these methods rely on bootstrap; a nonparametric technique that randomly re-samples the acquired data to generate new sets of data. This provides a distribution of measurements, which can be used to assess the statistical errors of a given metric with minimal assumptions about the noise in the original data. The basics of bootstrap applied to diffusion-weighted MRI are to acquire M datasets with the same parameters. Then, N new datasets are generated by randomly selecting subsets of data from each of the M original datasets (12). DTI reconstruction is then applied to each of the N bootstrap datasets, yielding N tensors with their derived metrics (e.g., fractional anisotropy, mean diffusivity, axial/radial diffusivities). Out of these N sets of processed bootstrap data, one is able to generate a distribution of a given metric and compute statistics on each of these distributions (e.g., mean, median, variance). If the quantity of interest is a vector (such as the principal diffusion direction), angular uncertainty can be assigned based on the confidence interval of the angular dispersion (7). Recent developments have extended the bootstrap method to estimate the uncertainty of the ODF first maximum (13–15), or used estimates from a Bayesian framework (16). While authors introduced original frameworks, these methods focus on the directionality of a single principal ODF maxima for direct application to probabilistic tractography.

Due to the relatively low number of original data used in repetition bootstrap (~5–10), there is a downward bias when estimating the angular uncertainty (17). To overcome this issue, Chung et al. have introduced the “repetition bootknife”, which combines repetition bootstrap with jackknife sampling (18). Given M repetitions, the key concept is to randomly omit one sample from the original dataset, and perform bootstrap on the remaining $(M-1)$ samples.

In the present study, we extended the work of bootstrap applied to HARDI with QBI reconstruction to provide a more generic measure of ODF reproducibility which, in addition to being useful for voxel-wise assessment of complex fiber regions, is potentially valuable for supporting HARDI-based metrics such as the Generalized Fractional Anisotropy (GFA) (2). The sensitivity and robustness of this method were evaluated using simulation studies where variable amount of noise was added to real data. Original contributions include 1) a confidence interval measure for secondary diffusion directions and 2) the study of the impact of acquisition parameters on the reproducibility of HARDI measurements, namely the echo time (TE) and the diffusion sensitization (b-values).

MATERIAL AND METHODS

MRI Acquisition

One healthy volunteer was scanned at 3 Tesla (TIM Trio, Siemens Medical Systems) using a 32-channel RF head coil (19). High angular resolution diffusion images were acquired using a twice refocusing EPI sequence with the following parameters: Axial orientation, Field of view = 207 mm, Number of slices = 18, covering the corpus callosum and aligned with the anterior commissure / posterior commissure (AC-PC), Slice thickness = 1.5 mm (no gap),

TR/TE = 3000/121 ms, In-plane resolution = $1.5 \times 1.5 \text{ mm}^2$, Parallel imaging (R = 2, GRAPPA reconstruction with 24 reference lines), Phase-encoding direction = antero-posterior, Bandwidth = 1510 Hz/pixel, echo spacing = 0.76 ms, b-value = 3000 s/mm², Number of diffusion-encoding directions = 60 (plus three b=0 images), Acquisition time = 3 min 12 s.

The acquisition was repeated 8 times, yielding a total acquisition time of ~25 min. The diffusion-gradient encoding set was built from the electrostatic repulsion algorithm that provided an evenly distributed diffusion-gradient scheme (20). Gradient vectors are listed in Table 1. Here we used 60 directions as a compromise between the minimum number of directions required to robustly reconstruct the Q-Ball ODF (21) and the total acquisition time acceptable for a human subject. Also, 60 directions were used by Whitcher *et al.* for performing wild bootstrap on Q-Ball ODF (22). In order to estimate subject motion we used interspersed b=0 images every 20 diffusion-weighted volumes. Therefore, three b=0 images were acquired per repetition (24 in total). No cardiac gating was used.

To demonstrate the efficiency of bootstrap metrics in assessing the quality of HARDI measurements, additional datasets were acquired with different imaging parameters. Since

the MR signal decreases by a factor of e^{-TE/T_2} , increasing the TE decreases the SNR of diffusion-weighted data and therefore the ability to reconstruct the ODF robustly. Another parameter affecting ODF reconstruction is the b-value (4–6). Increased b-value yields lower SNR but higher contrast-to-noise ratio for diffusion weighting. We therefore acquired two additional datasets with the exact same parameters and slice prescription as before, one with a TE of 121 ms and the other with a TE of 91 ms, both with a b-value of 1000 s/mm². Number of measurements was 6 for both additional datasets.

Optimum SNR coil combination—To study the relationship between reproducibility metrics and the thermal SNR in HARDI measurements, we acquired an SNR map at the same slice location used to generate bootstrap metrics. The SNR map was obtained by measuring the coil sensitivity and noise covariance to reconstruct the optimum SNR coil combination (23). The optimum SNR map was then corrected on a pixel-by-pixel basis to account for the SNR bias introduced by the magnitude detection. This correction for magnitude bias in array data was described by Constantinides *et al.* (24) and is dependent on both the SNR and the number of array channels. We extended the analysis of Constantinides *et al.* to include 32-channel data and applied the correction as a look-up table to the SNR maps.

Preprocessing

Motion Correction—One sensitive aspect of the long acquisition time imposed by regular bootstrap is subject motion. One has to correct for subject motion before processing the data, otherwise the measured reproducibility will likely reflect subject motion rather than other parameters of interest. Standard methods exist to register diffusion-weighted images on a target image, which could either be the averaged diffusion-weighted image or the b=0 image (25,26). However, these motion correction techniques are no longer robust when going to large b-value ($>3000 \text{ s/mm}^2$), where the SNR in individual diffusion-weighted images is lowered. This especially holds true in data acquired with small voxel size, as done in this study (1.5 mm isotropic).

For these reasons we relied on b=0 images regularly interspersed within each diffusion-weighted dataset in order to robustly estimate subject motion. Motion correction was conducted in two steps using FLIRT, part of the FSL package (27). The first step consisted of inter-repetition registration and the second step consisted of intra-repetition registration.

First, all $b=0$ images have been registered to the first $b=0$ image of the first repetition (rigid-body registration with 6 degrees of freedom). To limit the amount of image interpolations, only the transformation matrix has been saved and used in the following step. Second, each of the three $b=0$ images of a given repetition was co-registered to the first $b=0$ image of that repetition (rigid-body registration with 6 degrees of freedom). Each transformation matrix was then multiplied by the respective inter-repetition transformation matrix (translation matrices were added and rotation matrices were multiplied). The resulting transformation matrices were then applied to diffusion-weighted images.

Bootstrap Data Processing

Figure 1 shows a workflow of the overall bootstrap procedure with metric estimation. Here we used repetition bootknife and introduced two reproducibility metrics. One metric assessed the reproducibility of the overall shape of the ODF using the Jensen-Shannon divergence, while the other provided an angular confidence interval associated with the directions of the local maxima of the ODF; a role analogous to the angular confidence intervals of the DTI fiber direction in a conventional DTI bootstrap analysis.

The bootstrap implementation followed three main steps: 1) generation of bootstrap data, 2) reconstruction of Q-Ball ODFs and 3) estimation of bootstrap metrics. The method has been implemented in Matlab (The MathWorks, USA) using compiled C-functions for faster computation. On a 30×30 pixels ROI, it takes about 30 minutes to run the whole pipeline on a MacBook Pro, CPU 2.4 GHz with 4GB RAM. The developed software is platform-independent and is freely available upon request.

Generating Bootstrap Data—Out of the 8 original datasets, 500 bootstrap datasets were created by randomly picking a value among one of the original datasets using jackknife sampling. This was achieved on a voxel-by-voxel basis.

Processing Bootstrap Data—For each bootstrap dataset, Q-Ball diffusion ODFs were estimated using the method described in (6). The HARDI signal was expressed as a spherical harmonic (SH) series of order L and the Funk-Radon transform was solved using the Funk-Hecke theorem. The final ODF reconstruction, Ψ , in direction (θ, φ) can be written as

$$\Psi(\theta, \varphi) = \sum_{k=0}^L \sum_{m=-k}^k 2\pi P_k(0) c_k^m Y_k^m(\theta, \varphi) \quad [1]$$

where Y_k^m denote SH of order k and degree m , c_k^m are the SH coefficients describing the input HARDI signal and P_k is a Legendre polynomial of order k . We imposed a Laplace-Beltrami regularization criterion while estimating the SH coefficients c_k^m , as done in (28). In this study we used spherical harmonic decomposition of order 4 and a regularization parameter of 0.006 (28). In the literature the ODF has also been referred to as the “fiber ODF”, which is a sharp version of the diffusion ODF that could be obtained using deconvolution methods (29–31). In this study we only used the diffusion ODF, although the methodology could, in principle, be used with any other HARDI reconstruction technique. Therefore in the rest of the manuscript, the term ODF will refer to the diffusion ODF from QBI.

From the reconstructed ODF, we computed the generalized fractional anisotropy (GFA) (2), which is a HARDI anisotropy measure similar to the popular DTI fractional anisotropy (FA) (32). As an extension of the FA, the GFA is defined as the standard deviation divided by the

root mean square of the ODF. Hence, it is also a measurement of the anisotropy but it is generalized throughout more than three eigenvalues. Thus we have

$$\begin{aligned} GFA &= \sqrt{\frac{\sum_{i=1}^n (\Psi(u_i) - \langle \Psi \rangle)^2}{(n-1) \sum_{i=1}^n \Psi(u_i)^2}} \\ &= \sqrt{1 - \frac{(c_0^0)^2}{\sum_{k=0}^L \sum_{m=-k}^k (c_k^m)^2}} \end{aligned} \quad [2]$$

where

$$\langle \Psi \rangle = \left(\frac{1}{n} \right) \sum_{i=1}^n \Psi(u_i) \quad [3]$$

is the mean of the ODF.

Bootstrap Metrics

Jensen-Shannon divergence: In this paper we extended the bootstrap method from the evaluation of uncertainty in scalars (such as FA and eigenvalues) or vectors (such as the direction of the principal eigenvector), to comparing distributions such as the ODF itself. Since the ODF summarizes all physical findings of the single-shell HARDI diffusion measurement it is desirable to have a quantitative assessment of its reproducibility. In general, several distance measures have been used for comparing the distributions relevant to diffusion data. For example, the root mean sum-of-squares (6,28,33,34), the Kullback-Leibler divergence (2,4,35) or the Jensen-Shannon divergence (36) have been used to compare the ODFs generated from different reconstruction methods or to quantify fiber complexity. A recent study supported the use of Euclidean metrics to compute distances between diffusion tensors (37).

Here we used the Jensen-Shannon divergence (JSD) to quantify the similarity between the distribution of each bootstrap ODF and the distribution of the mean ODF. The mean ODF was generated by first converting the estimated Q-Ball ODF from the spherical harmonics into samples on a sphere. From the 60 direction measurements, we estimated our spherical functions with an SH series of order 4, which has 15 coefficients. Then, we can project these 15 SH basis coefficients to any number of spherical values on sphere. Here, we chose 724 spherical values as a compromise between the desired angular precision and computational time, in agreement with the literature (13,14,38). The 724 samples have been distributed equally on the sphere using the algorithm proposed in (39). Then, ODF from each bootstrap sample was averaged on a sample-by-sample basis. Note that this procedure was equivalent than computing the ODF from the averaged data, due to the linearity of the transformation.

Similarly to the root mean sum-of-square, the JSD measures the distance between two probability distributions. In this case, the probability distribution is the ODF as defined by discrete sampling on the sphere. The JSD is defined as:

$$JSD(P, Q) = \frac{D_{KL}(P, M) + D_{KL}(Q, M)}{2} \quad [4]$$

$$M(i) = \frac{P(i) + Q(i)}{2} \quad [5]$$

where $P(i)$ and $Q(i)$ are the distributions of the discrete ODF reconstructed from HARDI data, i is the index of sampling of the ODF (i.e., $i=1..724$), and $D_{KL}(\bullet, \bullet)$ is the Kullback-Leibler divergence, defined as:

$$D_{KL}(P, Q) = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right) \quad [6]$$

The notable advantage of using the JSD over the Kullback-Leibler divergence is that the former metric is symmetric, i.e., $JSD(P, Q) = JSD(Q, P)$.

Confidence interval: One goal of diffusion MRI is to map brain connections using fiber tractography. It is therefore desired to estimate the reliability of the reconstructed fiber pathways. This could be achieved by computing uncertainty measures on ODF maxima, similarly to what have been proposed for DTI (8). Only here, the difficulty was the potential presence of more than one diffusion direction, as opposed to a single principal direction in DTI. One original contribution of this paper was to derive maps of angular uncertainty, not only for the first maximum but also for the 2nd and eventually the 3rd maximum – providing they existed. In order to achieve this, one solution is to first sort the maxima of each bootstrap data based on their value on the ODF, and then compute the angular error on a rank-wise basis. Considering maxima of rank i , the angular error $\theta_{i,j}$ (in degree) would be computed between the maximum vector $\hat{v}_i$ of the mean ODF and the maximum vector $v_{i,j}$ of the j^{th} bootstrap ODF:

$$\theta_{i,j} = \frac{180}{\pi} \cdot \arccos(v_{i,j} \cdot \hat{v}_i) \quad [7]$$

However, achieving this is problematic since variability of the bootstrap samples can potentially lead to switching between ranks of ODF maxima. In the case where the first and second maxima are similar in size, variability would cause them to be ranked differently across bootstraps and the angular uncertainty of each maximum would be overestimated (just their labels would be reversed). Such situation would happen in regions containing crossing fibers with similar number and density. In the latter case, angular uncertainty would be overestimated – especially when fibers cross at $\sim 90^\circ$, as supported by our simulations (data not shown). To overcome this undesirable bias of the metric, we did not sort the maxima on each bootstrap ODF, but only on the mean ODF. Then, the rank of each maximum was defined so that the sum of the angles between each maximum of the mean ODF and each maximum of the bootstrap ODF was minimized. Hence, the rank was obtained by finding the permutation function f that minimized the sum of the dot product:

$$\min_f \sum_{k=1}^n |v_{f(k),j} \cdot \hat{v}_k| \quad [8]$$

where n is the number of maxima for the j^{th} bootstrap ODF. This approach certainly induces a downward correction in the estimation of the angular dispersion, but better supports the use of CI measures to derive cones of uncertainty for tractography applications (7) when all

maxima of the ODF – no matter what their ranks are – are used in probabilistic tractography. In that application the rank of extracted maxima is not pertinent. Of greater interest is to compute CI of all principal directions within each voxel, to be used as inputs for multiple-fiber probabilistic tractography algorithms. According to this rationale, the angular error $\theta_{i,j}$ was then defined as:

$$\theta_{i,j} = \frac{180}{\pi} \cdot \arccos(\tilde{v}_{i,j} \cdot \hat{v}_i) \quad [9]$$

where $\tilde{v}_{i,j} = v_{f(i),j}$ is the maximum vector of the j^{th} bootstrap ODF as defined by Equation [5]. Note that if $\tilde{v}_{i,j}$ did not exist, its value was omitted and the angular error was computed from a reduced number of bootstrap samples. A 95% CI was then calculated using the percentile method, as done for DTI (7). The i angular errors were ordered and the angle corresponding to 95% of the total number of samples were selected to obtain a statistical threshold of $P=0.05$. In the present study, CI maps were computed for the first and second maxima of the ODF.

Finding ODF Maxima—Extraction of ODF maxima was performed via an algorithm that finds local maxima on the discrete ODF. Local maxima were identified by testing each sample with its neighbors. The number of neighbors was set to $s/30$, s being the number of samples on the ODF. For a sampling of 724 points around the sphere, the number of neighbors was therefore 24. Here, Q-Ball ODFs were reconstructed using up to the 4th order of spherical harmonics, which produced relatively smooth ODF. Based on preliminary results, finding local maxima on smoothed ODF using an angular threshold of 12° (corresponding to 24 neighbors) has been found to robustly detect local maxima.

Due to the axial symmetry of the ODF, most maxima had their antipodal correspondent. This was taken into account to estimate the “true” number of maxima reflecting the number of crossing fibers at a given voxel. Hence, after finding local maxima, antipodal pairs were identified based on the angle between them ($\sim 180^\circ$). Following this step, the bisection between two antipodal maxima was defined as one maximum diffusion direction. Using this method with a spherical harmonic order of 4, each voxel usually produced one to three maxima. Among identified maxima, those within 45° of a local maximum were discarded, the local maximum being defined as having the highest value on the ODF. This was done to limit the impact of ODF peaks due to noise, given that spherical harmonic order of 4 does not allow to resolve crossing fibers at less than 65° (33). Also, the identified local maxima with magnitudes lower than 33% of the maximum ODF value were discarded, as done in (13). This threshold was empirically set to avoid picking maxima out of noise. Maxima extraction was performed voxel-wise on each bootstrap dataset and on the mean bootstrap dataset.

Maxima of the mean ODF were then sorted with respect to the probability of diffusion direction. This step consisted of labeling the first, second, third and subsequent diffusion directions, based on the likelihood of water diffusion given by the ODF. As mentioned previously, this ordering step was not required for computing CI maps. Therefore, it was not applied to each bootstrap ODF – only to the mean ODF. After maxima extraction, angular error was computed between the maxima of the mean ODF and the maxima of each bootstrap ODF. CI maps were then generated for the first and for the second maxima of the ODF.

Impact Of Noise

Variable levels of Rician noise were added to the original dataset (TE = 121 ms, b-value = 3000 s/mm²) to assess the sensitivity of the bootstrap metrics to image noise levels. First, SNR was measured on a single b=0 image by dividing the mean signal on a 10×10 region of interest (ROI) in the white matter, by the standard deviation of another 10×10 ROI in the background. The latter standard deviation was corrected for Rayleigh distribution, dividing

its value by $\sqrt{2 - \pi/2}$. The resulting SNR was 7.9. Note that due to the use of a multichannel array, this method did not properly estimate the *true* SNR. However in that case, the objective was only to get an approximate starting value used as a reference for the noise simulations. Then, fictive Rician distributed noise was added to the data with an amplitude of 50% 100% or 200% of the original background standard deviation as in (11). These datasets with increased noise levels were then analyzed by the bootstrap method.

Robustness Of The Method

Robustness of the bootstrap method was evaluated as a function of the number of original datasets and the number of bootstrap datasets generated. The analysis was rerun using between two and eight of the original datasets to generate 200 bootstraps. In a separate assessment, the eight original datasets were used to generate between 50 to 1000 bootstrap datasets. These evaluations were repeated for the various levels of added noise: 0%, 50%, 100% and 200%. The number of original or bootstrap data required to reach a steady state was evaluated by computing the root mean square error (RMSE) between the $i+1$ and the i iteration. This computation was done on a 30×30 axial ROI, positioned on the posterior-left region and covering part of the ventricle, grey matter and white matter with single and crossing fibers (see Figure 2b).

Impact Of Acquisition Parameters

The bootstrap metrics were evaluated for two acquisition parameters: TE and b-value. To evaluate the effect of TE, data were acquired at TE of 91 ms and 121 ms, both at b-value of 1000 s/mm². To evaluate the effect of b-value, data were acquired at b-value of 3000 s/mm² and 1000 s/mm², both at TE of 121 ms. Each comparison was done using six original datasets. Due to potential subject motion between datasets, a third step was added to the motion correction pipeline, which consisted of co-registering the three datasets together (using the first b=0 image in each dataset). JSD and CI maps were then estimated on a single slice for each dataset as described previously in the methods.

Note that we focused on a single slice rather than on a specific tract, because tractography was not the only application of our method, which could also be used voxel-wise for supporting the reliability of HARDI-based metrics in the context of clinical studies. Also we did not want to limit the applicability of our findings to a particular tractography method and therefore focused on the ODF (which forms an input to most tractography methods).

RESULTS

Q-Ball ODFs were successfully reconstructed and maxima extracted from each bootstrap dataset. To illustrate the variability of Q-Ball estimate across bootstraps, Figure 2 shows maps of ODFs and extracted maxima on a region that includes single and crossing fibers (see FA Colormap on Figure 2d for anatomical landmarks). Variability in both the shape and orientation of the ODF is noticeable across bootstrap (Figure 2c). Extracted maxima are also plotted for ten bootstrap samples and overlaid on the GFA map for the same region (Figure 2e). Large maxima dispersion is noticeable in isotropic regions such as in the ventricles and the grey matter. Dispersion of the first maximum (green) is the lowest in the corpus

callosum, with a less coherent dispersion of the second maximum (red) attributed to noise. In a region of crossing fibers, both the first and second maxima exhibit low dispersion.

Reproducibility Measurements

Figure 3b shows JSD map for an axial slice. The JSD is higher in the ventricles and in the inter-hemispheric fissure, as expected due to the isotropic nature of water diffusion in the cerebrospinal fluid. In the parenchyma, JSD values are the lowest with no clear distinction between the gray and the white matter. Interestingly, JSD values tend to increase toward the center of the brain, possibly reflecting the lower sensitivity of array coils in deep tissues, as illustrated by the map of thermal SNR (Figure 3c). Similarly mirroring the SNR map, patterns of high JSD value extend toward the frontal and sensori-motor regions of the left hemisphere. This comparison suggests a direct relationship between the reproducibility of ODF reconstruction and the SNR, a hypothesis further confirmed by the added-noise simulation studies. Maps of CI for the first and second maxima are shown in Figure 3e–f. The location of the white matter is clearly visible on the CI map of the first maxima, where 95% angular confidence is achieved with a CI of 0–10°, whereas angular confidence in the CSF or grey matter typically reach values of about 70–90°. In the CI map of the second maxima (where it exists), most values range from 70° to 90°, with no clear delineation of anatomical structure such as the cortical grey matter, which was expected to exhibit better confidence intervals in secondary directions when compared to purely isotropic regions such as the cerebrospinal fluid. However, reduced confidence interval of the second maxima is noticeable in known regions of crossing fibers, such as at the intersection between the corpus callosum and the posterior corona radiata (Figure 3g–i). The zoomed panel is a good representation of possible cases of single and crossing fibers. ODFs on the outer columns mostly contain the corpus callosum (left) and longitudinal tracts (right), therefore exhibit low uncertainty for the first maximum and either no (left) or very high uncertainty (right) for the second maximum. The middle column however exhibits ODFs with two distinct maxima, as reflected by the low uncertainty of the first and second maxima. Note that red lines (indicating the secondary maxima) exist for the voxel on the left bottom corner of Figure 3g while its corresponding voxel is represented by the black color on Figure 3i (indicating no secondary maximum). The reason is that the number of maxima attributed to each voxel was based on the mean ODF (averaged across 500 bootstraps), whereas the maxima that are shown on Figure 3g arise from 10 individual bootstrap samples. Therefore it is likely that for some bootstrap samples, a 2nd maximum was discovered, although there was no 2nd maximum on the mean ODF.

Impact Of Noise

Sensitivity of the bootstrap method was evaluated using realistic simulations with variable amount of added noise (0%, 50%, 100%) (Figure 4). In this figure, ODFs and maxima dispersion show the impact of noise on the Q-Ball reconstruction. Results of the simulation show a direct relationship between the level of SNR and JSD values, as suggested by the spatial similarity between the JSD and image SNR maps. Similarly, higher noise levels yield higher uncertainty in the CI maps of the first maxima.

Robustness Toward Bootstrap Parameters

Figure 5 shows the robustness of the bootstrap method as a function of the number of original data (2,3,6 or 8 original data sets and 500 bootstraps) for various levels of added noise. The region of interest is the same as that one shown in Figure 2b. Results suggest that JSD and CI estimation is inaccurate when insufficient number (<4) of original data is used. Figure 6 illustrates how the number of bootstraps (50 to 1000 from 8 original datasets) impacts JSD and CI estimates. Similarly to the previous simulation, estimation of JSD and CI is inaccurate with low number of bootstraps, although the differences between the results

with low or high number of bootstraps is relatively small. Figure 7 shows the RMSE between the i and the $i-1$ maps, where i is either the number of original datasets or the number of bootstraps. With more than 5 datasets, change in RMSE for JSD and CI exhibits a relatively small monotonic decay, meaning that there is no substantial gain in accuracy in the estimate of JSD and CI maps when adding more than 5 datasets. With >500 numbers of bootstraps, RMSE seems to reach a plateau, therefore suggesting that 500 bootstraps achieve a robust estimate of JSD and CI metrics.

Note that the trend observed on the CI is somewhat counterintuitive, as for a given number of dataset or for a given number of bootstrap, the RSME of CI decreases when the amount of noise increases. What is showed on this plot is not the actual CI value, but rather the difference in CI between two conditions (with an incremental dataset or number of bootstraps). Results suggest that for highly noisy data (e.g. 200%), the overall CI (i.e. CI averaged within an ROI encompassing white matter, gray matter and CSF) does not change much when additional data or bootstrap samples are provided. This is due to the fact that for noisy data, the CI is already high (>80°), no matter how many data or bootstrap samples are provided.

Impact Of TE

The sensitivity of the bootstrap method to pulse sequence parameters was evaluated by acquiring HARDI data at two different TEs (91 ms and 121 ms), where other parameters were kept the same (b-value was 1000 s/mm²). The SNR ratio between the two datasets could be evaluated using the following relationship:

$$\frac{SNR_1}{SNR_2} = e^{\frac{-(TE_1 - TE_2)}{T_2}} \quad [9]$$

Assuming a T_2 of 74 ms in the white matter at 3T (40), the TE=91 ms dataset has about 50% more SNR when compared to the TE=121 ms dataset. Due to this increase in SNR at the lower TE, the added-noise simulations suggest that improved reproducibility of ODF estimate is expected for the low TE dataset. Results of JSD and CI maps are shown in Figure 8 verifying a higher ODF reproducibility and lower angular uncertainty at TE=91 ms versus TE=121 ms.

To better characterize differences between the two TE measurements, a white matter mask was created from one dataset and used to generate white matter histograms of JSD and CI measures for both conditions (Figure 8c). The mask was generated using an intensity threshold of 0.12 on the GFA axial map from the TE=91 ms dataset. The JSD and CI histograms were distinct for the two different TE values as determined by a 2-tailed, 160 degrees of freedom Student T-test performed on the assumed Gaussian data. Results of the T-test support unequal means between the two distributions for the JSD ($T=9.66$, $P<10^{-12}$) and for the CI ($T=6.08$, $P<10^{-8}$). Although these distributions exhibit a slight positive skew, Gaussian fit was used in order to perform statistical comparison between means, as well as to compare the shape of either distribution.

Impact Of B-value

Figure 9 shows analysis of the JSD and CI metrics applied to HARDI data acquired at two different b-values (3000 s/mm² and 1000 s/mm²). Due to higher SNR of diffusion-weighted images at lower b-value, reproducibility of ODF estimate was expected to be higher. This hypothesis is verified in the JSD map, which exhibits a higher overall reproducibility at

b=1000 versus b=3000. The contrast of the GFA map is also higher at low b-value, providing better delineation of the white matter structure.

A trade-off for using low b-value however was the lower angular resolution for detecting multiple fiber orientations. Again, this hypothesis is verified in the CI maps, which exhibits lower uncertainty in white matter regions at higher b-value. This observation is especially marked in regions of crossing fibers, as illustrated by the zoomed panel of Figure 9, which focuses on the intersection of the corpus callosum, the anterior corona radiata and longitudinal tracts (black circle). Interestingly, CI values in gray matter and CSF are higher at b=3000 s/mm². Thus, the higher contrast in CI value between white matter and non-white matter at high b-value suggests an improved ability for probabilistic tractography. This result highlights the relevance of CI maps both inside and outside the white matter for assessing acquisition parameters.

The white matter mask of Figure 8 was used to compute histograms of JSD and CI for both b-values (Figure 9c). Histograms were fitted to a Gaussian distribution and Student T-test was performed (2-tailed, 190 degrees of freedom). Results of the T-test support unequal means between the two distributions for the JSD ($T=9.94$, $P<10^{-12}$) and for the CI ($T=2.62$, $P<0.01$). Unlike the TE case, the shape (i.e., standard deviation of the fitted Gaussian function) of JSD histograms is very similar between the b-value conditions and only the mean was significantly upward shifted for the high b-value condition. This observation suggests that the shape of the JSD histogram may reflect differences in the biophysical properties of the measurement (beyond simple SNR).

DISCUSSION

In this paper we presented a repetition bootstrap method to assess the quality of HARDI data. The proposed metrics targeted two different levels of quality assurance: the JSD metric measures the reproducibility of ODF shape and the angular CI measures the uncertainty of ODF maxima. Results from brain data showed a direct relationship between bootstrap metrics and thermal SNR. Acquisitions at variable TEs and b-values demonstrated the efficiency of the proposed method for characterizing the reproducibility and accuracy of the ODF and its derived maxima. This section will first address some key features of the bootstrap method, including its robustness and limitations. Then, a possible extension of the method is proposed regarding the type of bootstrap method. Finally, various applications of the method are proposed.

Bootstrap Metrics

Impact Of Noise—Evaluation of bootstrap metrics was mainly focused on the impact of noise as assessed by adding noise to the data, by acquiring data at two different TEs and by comparing JSD and SNR maps. In the latter case, we found that JSD maps reproduced the approximate spatial pattern of the sensitivity profile of the 32-channel array coil used here, i.e., lower sensitivity in central regions relative to peripheral regions (19). This finding is supported by previous studies that aimed at measuring the reproducibility of diffusion-weighted measurements in DTI using multichannel array coils (37,41). Similar findings were observed using the Kullback-Leibler divergence to measure the accuracy of the Q-Ball ODF at various levels of noise (38).

Another interesting finding is that JSD histograms exhibited very different shape between acquisitions at different TEs and b-values. Comparing two different TEs, the mean and standard deviation of the low-TE histogram were lower (Figure 8c). However, comparing two different b-values, the mean was decreased (lower b-value) but the standard deviation was approximately unchanged (Figure 9c). In both cases (longer TE and higher b value) the

noise was increased. These changes in histogram shape highlighted differences in ODF reproducibility derived from differences in the acquisition beyond just the SNR; for example the reproducibility associated with forming an ODF from the different biophysical sensitivity associated with different b-value acquisitions. It should be mentioned that T2 relaxation in the white matter can significantly vary across regions (42). Therefore, a change in the TE had a different impact on the SNR throughout the whole brain.

Noise also increased the angular uncertainty of ODF maxima as measured in the CI maps when noise was added or when the TE was increased (producing noisier images). When the b-value was increased from 1000 to 3000 however, the CI histograms and maps were improved (*i.e.*, lower angular uncertainty), even though the higher b-value produced noisier images. Measuring CI of FA in DTI, Heim *et al.* found similar differences between histogram distributions at various levels of noise and smoothing (11).

B-value—The achievable angular resolution in HARDI depends both on the diffusion gradient scheme and the b-value. While other studies suggested the use of large b-value for better accuracy in depicting diffusion maxima (2,5), the present study demonstrated that larger b-value also provides better angular confidence interval of the ODF maxima (Figure 9). One side effect of using higher b-value is that it increases the minimally achievable TE, therefore further reducing the SNR. Metrics provided in this study may help optimizing the trade-off between the TE and the b-value.

Secondary Maxima—One of the original contributions in this study was the introduction of confidence interval maps of secondary diffusion directions. While CI maps of first maxima provided very consistent white matter maps, the relevance of CI maps of second maxima was more difficult to assess. Most values exhibited very large confidence intervals ($>70^\circ$) with no clear delineation of anatomical structure. However, a better confidence interval was noticeable in known regions of crossing fibers, such as at the intersection between the corpus callosum and the posterior corona radiata (Figure 3i). These findings suggest that the CI of the second maxima could be derived from HARDI measurements, providing consistent measure of uncertainty in regions with more than one principal diffusion direction.

CI maps of secondary maxima may find applications in multiple-fiber probabilistic tractography (13,14,16). Although they would not dramatically change the outcome of tractography results in dominant white matter pathways, CI measurements of secondary maxima may be of interest in regions with higher complexity. Future investigations are needed to evaluate the pertinence of this index using other reconstruction methods that provide sharper versions of the ODF, therefore better discrepancies between maxima.

Rotation of the B-matrix—Rigid registration in multiple directions diffusion-weighted imaging should be, in general, accompanied with an appropriate rotation of the B-matrix (43). For the within-dataset case, motion correction was performed between repetitions (*i.e.*, between the first and the eighth repetitions). Therefore, we could not easily have applied a rotation of the B-matrix as for a given voxel, the data were randomly picked from one of the eight repetitions, and a single B-matrix was then applied to the whole volume. In other terms, each voxel of the bootstrap-generated volume experienced a different motion, making the B-matrix correction approach less easy to apply. Although we acknowledge the fact that motion might have biased the estimation of ODF, the amount of rotations as assessed by the motion correction parameters given by FLIRT was marginal (less than 2°).

For the between-dataset case, the registration was performed between one dataset (e.g. b=1000, TE=91ms) and another dataset (e.g. b=3000, TE=91). Therefore, correcting for the

global registration by applying the same rotation to each row of the b-matrix wouldn't have changed our results, as the present study focused on the individual voxel, and not on the neighboring (e.g., tractography). Therefore, the derived JSD and CI metrics were not affected by potential rotations between datasets.

How Much Data for Bootstrap?

We evaluated the number of datasets needed to obtain reliable estimate of bootstrap metrics. Given the potential dependence of this parameter on the SNR of the measured data, we used simulations with variable amount of noise. Results suggest that a minimum of five datasets is required to reach robust estimate of reproducibility metrics. Similarly, for DTI using regular bootstrap, O'Gorman and Jones suggested at least five repetitions in order to have a robust estimate of reproducibility metrics (44). However it should be noted that the number of directions in HARDI acquisition impacts the quality of the ODF reconstruction for each bootstrap sample (35). O'Gorman and Jones investigated the impact of the acquisition scheme for DTI reconstruction (44). For 8, 16, 32 and 64-direction schemes, they found the minimum number of repetitions required to obtain robust estimate of the standard deviation of the trace of the tensor to be about five. Keeping in mind that HARDI Q-Ball and DTI are very different reconstruction techniques and that bootstrap-based evaluation of reproducibility underlies different assumptions about the noise, results of our simulations and those from O'Gorman and Jones therefore suggest that the proposed bootstrap method requires at least five repetitions to ensure accuracy in computation of bootstrap metrics. Regarding the number of bootstraps required to derive reliable metrics, our results suggest to use at least 500 bootstraps. In contrast, 100–200 bootstraps were suggested for DTI (11,44).

Here we used 60 directions for our HARDI acquisition protocol. We expect a bias in the estimate of the reproducibility of the Q-Ball reconstruction due to the selection of the diffusion-encoding set, given that increasing the number of directions will increase the accuracy of the ODF reconstruction (35), with a plateau after ~100 directions (45). Also, previous studies showed how the number of diffusion-encoding directions affect bootstrap estimates reproducibility (18,34). Therefore, any comparison should be conducted using the same encoding scheme.

Previous studies have also shown that angular uncertainty and ODF reproducibility depend on the b-value (38) as well as on the SNR (11,13,14,41). These parameters are therefore expected to significantly impact the required number of data for bootstrap (34).

Bootstrap Methods

In this study we used regular bootstrap with jackknife sampling. We showed that at least 5 repeated measurements were required, leading to 16 min of acquisition for covering a portion of the brain. If covering the whole brain (usually 60–70 slices for a thickness of 2 mm), TR would be at least much longer and the scan time would be at least three times longer. Unless fast imaging technique is used, such as compressed sensing or simultaneous multislice imaging (46), it is hard to apply the repetition bootstrap for real human data in a routinely basis. In contrast, model-based bootstrap techniques, such as residual bootstrap (13,14,18) and wild bootstrap (22,25,47), can provide equivalently robust estimations of measurement uncertainty as those from repetition bootstrap while only one complete DTI/HARDI data is normally required.

One advantage of regular bootstrap is its applicability to model-free approaches such as diffusion spectrum imaging (DSI). Also, acquiring independent HARDI measurements without any noise assumption potentially provides a less biased measure of the reproducibility of HARDI measurements. It should be mentioned however that previous

studies showed similar results between residual and repetition bootstrap (14,18,22,34,47). Lastly, it should be noted that the metrics proposed in this study could also be derived using residual or wild bootstrap techniques.

Possible Applications

As researchers aim to push the limits of diffusion-weighted imaging with higher b-values and smaller voxels, there is an increasing need to evaluate the reproducibility of these measurements. Recently, the proposed method has been applied to evaluate the benefit of a head gradient for diffusion imaging (48). The lower JSD reflected the higher SNR achievable by lowering the TE while keeping the same b-value. The proposed method has also been applied to evaluate a new diffusion pulse sequence based on simultaneous multislice imaging (46). By providing quantitative metrics to evaluate the reproducibility of the ODF, the authors demonstrated the equivalence in the quality of HARDI data, with and without enabling the multislice acquisition.

A second application domain for the proposed method concerns ODF reconstruction techniques. As previously mentioned, the JSD and CI metrics only need the discrete ODF to be generated, therefore other methods than QBI could be evaluated and compared for their sensitivity to noise and ability to produce maxima with low angular uncertainty. The measure of angular uncertainty derived from the bootstrap method could also be used as input for HARDI-based probabilistic tractography (13,31). Note that the reduced coverage used here might be limited for some tractography applications. The main reason for not covering the entire brain was to reduce the acquisition time to a reasonable 3 minutes per dataset, allowing multiple acquisitions possible for the bootstrap validation. Moreover the goal here was to demonstrate the utility of a bootstrap method to assess the reproducibility of the ODF. Researchers can adapt the proposed protocol by modifying the coverage according to their needs.

Finally, clinical applications will benefit from a statistical assessment on the reliability of HARDI datasets. Indeed, ODF-based metrics such as the GFA not only depend on the integrity of the white matter but also on the SNR and diffusion-gradient scheme (b-value and number of directions) of diffusion-weighted data (7,25). Therefore, reporting indices of reliability along with quantitative metrics is a possible avenue for researchers to gain confidence in HARDI studies. Also, indices of reproducibility can be used as quality insurance in longitudinal or multisite studies and to guide power analysis in clinical trials.

In conclusion, we presented a bootstrap method to assess the reproducibility of the ODF and angular uncertainty of the first and second ODF maxima derived from HARDI data. Quality assurance metrics were derived from the orientation distribution function, and proved to be efficient in relating bootstrap metrics with the SNR and b-value of HARDI data. This method may find applications in evaluating new hardware and pulse sequence in diffusion MRI, comparing ODF reconstruction techniques, developing probabilistic tractography and reporting reliability indices on quantitative metrics used for clinical studies.

Acknowledgments

Grant Support

National Institute of Health (P41RR14075, U01MH093765), Association pour la Recherche sur la Sclérose en Plaques (ARSEP).

REFERENCES

1. Le Bihan D, Mangin JF, Poupon C, et al. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging*. 2001; 13(4):534–546. [PubMed: 11276097]
2. Tuch DS. Q-ball imaging. *Magn Reson Med*. 2004; 52(6):1358–1372. [PubMed: 15562495]
3. Cohen-Adad J, Descoteaux M, Rossignol S, Hoge RD, Deriche R, Benali H. Detection of multiple pathways in the spinal cord using q-ball imaging. *Neuroimage*. 2008; 42(2):739–749. [PubMed: 18562214]
4. Canales-Rodriguez EJ, Melie-Garcia L, Iturria-Medina Y. Mathematical description of q-space in spherical coordinates: exact q-ball imaging. *Magn Reson Med*. 2009; 61(6):1350–1367. [PubMed: 19319889]
5. Cho K-H, Yeh C-H, Tournier J-D, Chao Y-P, Chen J-H, Lin C-P. Evaluation of the accuracy and angular resolution of q-ball imaging. *Neuroimage*. 2008; 42(1):262–271. [PubMed: 18502152]
6. Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Regularized, fast, and robust analytical Q-ball imaging. *Magn Reson Med*. 2007; 58(3):497–510. [PubMed: 17763358]
7. Jones DK. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magn Reson Med*. 2003; 49(1):7–12. [PubMed: 12509814]
8. Jones DK, Pierpaoli C. Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach. *Magn Reson Med*. 2005; 53(5):1143–1149. [PubMed: 15844149]
9. Lazar M, Alexander AL. Bootstrap white matter tractography (BOOT-TRAC). *Neuroimage*. 2005; 24(2):524–532. [PubMed: 15627594]
10. Hasan KM. A framework for quality control and parameter optimization in diffusion tensor imaging: theoretical analysis and validation. *Magn Reson Imaging*. 2007; 25(8):1196–1202. [PubMed: 17442523]
11. Heim S, Hahn K, Samann PG, Fahrmeir L, Auer DP. Assessing DTI data quality using bootstrap analysis. *Magn Reson Med*. 2004; 52(3):582–589. [PubMed: 15334578]
12. Pajevic S, Basser PJ. Parametric and non-parametric statistical analysis of DT-MRI data. *J Magn Reson*. 2003; 161(1):1–14. [PubMed: 12660106]
13. Berman JI, Chung S, Mukherjee P, Hess CP, Han ET, Henry RG. Probabilistic streamline q-ball tractography using the residual bootstrap. *Neuroimage*. 2008; 39(1):215–222. [PubMed: 17911030]
14. Haroon HA, Morris DM, Embleton KV, Alexander DC, Parker GJM. Using the model-based residual bootstrap to quantify uncertainty in fiber orientations from Q-ball analysis. *IEEE Trans Med Imaging*. 2009; 28(4):535–550. [PubMed: 19272997]
15. Jeurissen, B.; Leemans, A.; Tournier, JD.; Sijbers, J. Estimation of Uncertainty in Constrained Spherical Deconvolution Fiber Orientations; Proceedings of the 5th IEEE International Symposium on Biomedical Imaging; Paris, France. 2008. p. 907-910.
16. Fonteijn HM, Verstraten FA, Norris DG. Probabilistic inference on Q-ball imaging data. *IEEE Trans Med Imaging*. 2007; 26(11):1515–1524. [PubMed: 18041266]
17. Berman, JI.; Chung, S.; Hess, CP.; Mukherjee, P.; Lu, Y.; Henry, RG. Model-Based Bootstrap Resampling Methods for HARDI: Quantitation of ODF Uncertainty Without Multiple Acquisitions; Proceedings of the 15th Annual Meeting of ISMRM; Berlin, Germany. 2007. p. 1471
18. Chung S, Lu Y, Henry RG. Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. *Neuroimage*. 2006; 33(2):531–541. [PubMed: 16938472]
19. Wiggins GC, Triantafyllou C, Potthast A, Reykowski A, Nittka M, Wald LL. 32-channel 3 Tesla receive-only phased-array head coil with soccer-ball element geometry. *Magn Reson Med*. 2006; 56(1):216–223. [PubMed: 16767762]
20. Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn Reson Med*. 1999; 42(3):515–525. [PubMed: 10467296]
21. Alexander DC. Multiple-fiber reconstruction algorithms for diffusion MRI. *Ann N Y Acad Sci*. 2005; 1064:113–133. [PubMed: 16394152]

22. Whitcher B, Tuch DS, Wisco JJ, Sorensen AG, Wang L. Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging. *Hum Brain Mapp.* 2008; 29(3):346–362. [PubMed: 17455199]
23. Roemer PB, Edelstein WA, Hayes CE, Souza SP, Mueller OM. The NMR phased array. *Magn Reson Med.* 1990; 16(2):192–225. [PubMed: 2266841]
24. Constantinides CD, Atalar E, McVeigh ER. Signal-to-noise measurements in magnitude images from NMR phased arrays. *Magn Reson Med.* 1997; 38(5):852–857. [PubMed: 9358462]
25. Zhu T, Liu X, Gaugh MD, et al. Evaluation of measurement uncertainties in human diffusion tensor imaging (DTI)-derived parameters and optimization of clinical DTI protocols with a wild bootstrap analysis. *J Magn Reson Imaging.* 2009; 29(2):422–435. [PubMed: 19161198]
26. Sakaie KE, Lowe MJ. Quantitative assessment of motion correction for high angular resolution diffusion imaging. *Magn Reson Imaging.* 2010; 28(2):290–296. [PubMed: 19695824]
27. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage.* 2004; 23 Suppl 1:S208–S219. [PubMed: 15501092]
28. Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Apparent diffusion coefficients from high angular resolution diffusion imaging: estimation and applications. *Magn Reson Med.* 2006; 56(2):395–410. [PubMed: 16802316]
29. Jian B, Vemuri BC. A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. *IEEE Trans Med Imaging.* 2007; 26(11):1464–1471. [PubMed: 18041262]
30. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage.* 2007; 35(4):1459–1472. [PubMed: 17379540]
31. Descoteaux M, Deriche R, Knösche TR, Anwander A. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans Med Imaging.* 2009; 28(2):269–286. [PubMed: 19188114]
32. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med.* 1996; 36(6):893–906. [PubMed: 8946355]
33. Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magn Reson Med.* 2006; 56(1):104–117. [PubMed: 16755539]
34. Zhu T, Liu X, Connelly PR, Zhong J. An optimized wild bootstrap method for evaluation of measurement uncertainties of DTI-derived parameters in human brain. *Neuroimage.* 2008; 40(3):1144–1156. [PubMed: 18302985]
35. Zhan L, Leow AD, Jahanshad N, et al. How does angular resolution affect diffusion imaging measures? *Neuroimage.* 2010; 49(2):1357–1371. [PubMed: 19819339]
36. Chiang MC, Barysheva M, Lee AD, et al. Brain fiber architecture, genetics, and intelligence: a high angular resolution diffusion imaging (HARDI) study. *Med Image Comput Comput Assist Interv.* 2008; 11(Pt 1):1060–1067. [PubMed: 18979850]
37. Pasternak O, Sochen N, Basser PJ. The effect of metric selection on the analysis of diffusion tensor MRI data. *Neuroimage.* 2010; 49(3):2190–2204. [PubMed: 19879947]
38. Zhan L, Leow AD, Barysheva M, et al. Investigating the uncertainty in multi-fiber estimation in High Angular Resolution Diffusion Imaging. *Med Image Comput Comput Assist Interv.* 2009; 12:256–267.
39. Saff E, Kuijlaars A. Distributing many points on a sphere. *Math Intell.* 1997; 19(1):5–11.
40. Cox, EF.; Gowland, PA. Measuring T and T' in the brain at 1.5T, 3T and 7T using a hybrid gradient echo-spin echo sequence and EPI; Proceedings of the 16th Annual Meeting of ISMRM; Toronto, Canada. 2008. p. 1411
41. Farrell JAD, Landman BA, Jones CK, et al. Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. *J Magn Reson Imaging.* 2007; 26(3):756–767. [PubMed: 17729339]

42. Wansapura JP, Holland SK, Dunn RS, Ball WS. NMR relaxation times in the human brain at 3.0 tesla. *Journal of magnetic resonance imaging : JMRI*. 1999; 9(4):531–538. [PubMed: 10232510]
43. Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med*. 2009; 61(6):1336–1349. [PubMed: 19319973]
44. O'Gorman RL, Jones DK. Just how much data need to be collected for reliable bootstrap DT-MRI? *Magn Reson Med*. 2006; 56(4):884–890. [PubMed: 16958073]
45. Poupon C, Roche A, Dubois J, Mangin J-F, Poupon F. Real-time MR diffusion tensor and Q-ball imaging using Kalman filtering. *Med Image Anal*. 2008; 12(5):527–534. [PubMed: 18664412]
46. Setsompop, K.; Cohen-Adad, J.; McNab, J.; Gagoski, BA.; Wedeen, VJ.; Wald, LL. Improving SNR per unit time in Diffusion Imaging using a blipped-CAIPIRINHA simultaneous multi-slice EPI acquisition; Proceedings of the 18th Annual Meeting of ISMRM; Stockholm, Sweden. 2010. p. 187
47. Jones DK. Tractography gone wild: probabilistic fibre tracking using the wild bootstrap with diffusion tensor MRI. *IEEE transactions on medical imaging*. 2008; 27(9):1268–1274. [PubMed: 18779066]
48. Cohen-Adad, J.; McNab, JA.; Benner, T., et al. Improving high-resolution Q-Ball imaging with a head insert gradient: Bootstrap and SNR analysis; Proceedings of the 18th Annual Meeting of ISMRM; Stockholm, Sweden. 2010. p. 1606

Figure 1.

Pipeline of the bootstrap method. M diffusion-weighted datasets are acquired. N new datasets are then generated by randomly withdrawing samples out of the $M-1$ datasets on a voxel-by-voxel basis (jackknife sampling). Then, Q-Ball ODFs are estimated for each bootstrap dataset, allowing to compute ODF-based metrics with standard deviation across bootstrap samples. To obtain a measure of angular uncertainty on principal diffusion directions, ODF maxima are extracted and 95% confidence interval (CI) maps are computed.

Figure 2.

Variability of Q-Ball ODF across bootstrap samples. Axial GFA map (a) with zoomed panel overlaid ODF (b) shows the consistency of the reconstructed mean ODF, according to the known anatomical features as described by the FA colormap (d). Example of individual bootstrap ODFs (c) shows the variability of the ODF shape and orientation. Example of extracted maxima (e) illustrates the angular dispersion as quantified by the confidence interval maps. In this example, only the maxima of the first 10 bootstrap ODFs are shown for clarity purpose. Maxima are color-coded in respect to their rank, i.e., 1st (green), 2nd (red) or 3rd (blue) maxima. Subject orientation is indicated in the GFA map as anterior (A), Posterior (P), Left (L) and right (R). Labeled tracts are the corpus callosum (cc), inferior fronto-occipital fasciculus (ifo), inferior longitudinal fasciculus (ilf), posterior corona radiata (pcr), posterior thalamic radiation (ptr) and tapetum (tap).

Figure 3.

Reproducibility metrics. GFA (a), JSD (b), thermal SNR (c), number of maxima (d), CI of the first maximum (e), CI of the second maximum (f), mean ODFs showing the first 10 bootstrap maxima on the region circled in the CI maps (g), CI of the first (h) and second (i) maximum corresponding to the displayed ODFs. On the CI map of the second maximum, black color means that no value exists, i.e., only one maximum was found on these voxels. Imaging parameters for this dataset was $b=1000 \text{ s/mm}^2$ and $TE=121\text{ms}$.

Figure 4.

Impact of noise. Comparison of reproducibility metrics at various levels of noise added to the original data: 0%, 50% and 100%. From left to right column: GFA, ODFs overlaid on the GFA (ODF have been computed from a single bootstrap sample to emphasize the effect of noise), maxima dispersion showing first and second maxima of the ODF (only 10 bootstrap samples shown here), JSD and CI of the first maximum.

Figure 5.

Robustness toward number of original data with variable levels of added noise. JSD map (**a**) with 0%, 50% and 100% of added noise. CI of the first maximum (**b**) with 0%, 50% and 100% of added noise. Number of original data vary from 2 to 8. The selected region of 30×30 voxels corresponds to that one used in Figure 2b. Color scaling has been kept the same to allow direct comparison. Underestimation of JSD and overestimation of CI occur with insufficient number of original data.

Figure 6.

Robustness toward number of bootstrap (BS) data with variable levels of added noise. JSD map (a) with 0%, 50% and 100% of added noise. CI of the first maximum (b) with 0%, 50% and 100% of added noise. Number of bootstraps vary from 50 to 1000. The selected region of 30×30 voxels corresponds to that one used in Figure 2b. Similarly to the previous simulation, underestimation of JSD and overestimation of CI occur with insufficient number of bootstraps.

Figure 7.

RMSE of JSD and CI maps between two consecutive varying conditions, i.e., number of original data (top) and number of bootstrap data (bottom). Three different levels of noise have been added to the original data. We used a 30×30 axial ROI, positioned on the posterior-left region and covering part of the ventricle, grey matter and white matter with single and crossing fibers (see Figure 2b). The motivation for including non-white matter structures was to evaluate the reproducibility of the bootstrap methods in various structures. For each point i of the abscissa, the RMSE represents the distance between the i and the $i-1$ iteration. This figure suggests that for all levels of noise, about 5 repetitions and 500 bootstraps are necessary to reach a robust estimate of JSD and CI metrics.

Figure 8.

Evaluation of the bootstrap method at different TE. GFA, JSD and CI of the first maximum are shown for TE of 91 ms (**a**) and 121 ms (**b**). JSD maps show better reproducibility of ODF estimate at lower TE, due to higher SNR in diffusion-weighted measurements. Similar conclusion if drawn in CI maps, which overall exhibit lower angular uncertainty at TE=91 ms. For more thorough evaluation, histograms of JSD and CI metrics have been computed within a white matter mask (**c**). Number of bins used to generate both histograms was 100. As observed in either maps, more voxels exhibit low JSD values and low angular uncertainty at TE=91ms.

Figure 9.

Evaluation of the bootstrap method at different b-values. GFA, JSD and CI of the first maximum are shown for b-value of 3000 s/mm² (**a**) and 1000 s/mm² (**b**). JSD maps show better reproducibility of ODF estimate at lower b-value, due to higher SNR in diffusion-weighted measurements. However, CI maps show lower uncertainty at b=3000 s/mm², supporting the larger angular specificity of HARDI measurements at high b-value. For more thorough evaluation, histograms of JSD and CI metrics have been computed within the same white matter mask as in Figure 8 (**c**). As observed qualitatively, lower JSD values are found at lower b-value, and slightly lower CI values are found at higher b-value. Interestingly, both JSD histograms at low and high b-value exhibit similar shape (i.e., standard deviation).

Table 1

List of diffusion gradient vectors applied for each of the 63 acquired volume, consisting of three b=0 images and 60 diffusion-weighted volumes.

#	x	y	z
1	0	0	0
2	1.000	0.000	0.000
3	0.588	-0.381	0.713
4	0.896	0.018	-0.443
5	0.730	-0.416	-0.542
6	-0.154	0.965	0.210
7	0.117	0.777	0.618
8	-0.408	-0.420	0.810
9	0.211	0.758	-0.617
10	0.095	-0.971	-0.220
11	-0.136	-0.029	0.990
12	-0.143	-0.755	0.640
13	-0.635	0.749	0.189
14	0.081	0.973	-0.216
15	0.107	-0.026	-0.994
16	0.300	0.392	-0.870
17	-0.207	-0.807	-0.552
18	0.332	-0.036	0.942
19	-0.804	-0.592	0.058
20	0.337	-0.754	0.564
21	0	0	0
22	0.591	-0.764	-0.257
23	-0.089	-0.471	-0.877
24	-0.529	-0.437	-0.728
25	-0.368	-0.927	-0.070
26	0.311	0.928	0.206
27	-0.685	0.444	0.578
28	-0.030	-0.968	0.249
29	-0.337	0.749	0.571
30	-0.577	0.030	0.816
31	0.271	-0.752	-0.600
32	-0.936	0.190	-0.296
33	0.889	0.315	0.333
34	-0.886	0.039	0.462
35	0.545	0.591	0.595
36	-0.352	-0.021	-0.936
37	0.415	-0.902	0.121
38	-0.735	-0.002	-0.678
39	-0.739	0.615	-0.276

#	x	y	z
40	-0.257	0.412	0.874
41	0	0	0
42	-0.918	0.377	0.125
43	-0.258	0.770	-0.584
44	-0.591	0.439	-0.677
45	0.892	-0.191	0.410
46	0.681	0.148	0.718
47	-0.880	-0.315	-0.356
48	-0.621	-0.701	-0.351
49	0.895	-0.435	-0.096
50	-0.164	0.423	-0.891
51	-0.507	-0.781	0.364
52	0.107	-0.452	0.886
53	0.883	0.439	-0.167
54	0.600	0.025	-0.799
55	0.694	0.701	0.162
56	-0.393	0.902	-0.177
57	-0.762	-0.408	0.504
58	0.211	0.402	0.891
59	0.377	-0.385	-0.843
60	0.735	-0.615	0.286
61	-0.987	-0.150	0.059
62	0.528	0.810	-0.254
63	0.664	0.477	-0.576