Journal of Magnetic Resonance 305 (2019) 22–30

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

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

Improved accuracy of apparent diffusion coefficient quantification using a fully automatic noise bias compensation method: Preliminary evaluation in prostate diffusion weighted imaging

Xiaodong Zhong [a,*], Brian M. Dale [b], Marcel D. Nickel [c], Stephan A.R. Kannengiesser [c], Berthold Kiefer [c], Mustafa Bashir [d,e]

[a] MR R&D Collaborations, Siemens Healthcare, Los Angeles, CA, United States
[b] MR R&D Collaborations, Siemens Healthcare, Cary, NC, United States
[c] MR Application Development, Siemens Healthcare GmbH, Erlangen, Germany
[d] Department of Radiology, Duke University Medical Center, Durham, NC, United States
[e] Center for Advanced Magnetic Resonance Development, Duke University Medical Center, Durham, NC, United States

ARTICLE INFO

Article history:
Received 15 April 2019
Revised 11 May 2019
Accepted 20 May 2019
Available online 21 May 2019

Keywords:
Diffusion
ADC
Noise bias
Quantification
Prostate
Endorectal coil

ABSTRACT

Noise in diffusion magnetic resonance imaging can introduce bias in apparent diffusion coefficient (ADC) quantification. Previous studies proposed methods that are site-specific techniques as research tools with limited availability and typically require manual intervention, not completely ready to use in the clinical environment. The purpose of this study was to develop a fully automatic computational method to correct noise bias in ADC quantification and perform a preliminary evaluation in the clinical prostate diffusion weighted imaging (DWI). Using a pseudo replica approach for the noise map calculation as well as a direct mapping and a stepwise Chebychev polynomial modelling approach for the ADC fitting, a fully automatic noise-bias-compensated ADC calculation method was proposed and implemented both on the scanner and offline. The proposed method was validated in a computer simulation and a standardized diffusion phantom with ground-truth values. Two in vivo studies were performed to evaluate the proposed method in the clinical environment. The first in vivo study performed acquisitions using a clinically routine prostate DWI protocol on 29 subjects to evaluate the consistency between simulated and empirical results. In the second in vivo study, prostate ADC values of 14 subjects were compared between data acquired with external coils only and reconstructed with the proposed method vs. acquired with external combined with endorectal coils and reconstructed with the conventional method. In statistical analyses, $p < 0.05$ was regarded as significantly different. In the computer simulation, the proposed method showed smaller error percentage than the other methods and was significantly different ($p < 2.2 \times 10^{-16}$). With low signal-to-noise ratio (SNR), the conventional method underestimated ADC values compared to the ground truth values of the diffusion phantom, while the results of the proposed method were more consistent with the ground truth values. Statistical analyses showed no significant differences between measured and simulated results in the first in vivo study ($p = 0.5618$). Data from the second in vivo study showed that agreement between ADC measured with external coils only and combined coils was improved for the proposed method (mean bias: 0.04×10^{-3} mm^2/s, 95% confidence interval (CI) = $[-0.01, 0.09] \times 10^{-3}$ mm^2/s, $p = 0.187$), compared to the conventional method (mean bias: -0.12×10^{-3} mm^2/s, 95% CI = $[-0.17, -0.06] \times 10^{-3}$ mm^2/s, $p < 0.0001$). The proposed method compensates noise bias in low-SNR diffusion-weighted acquisitions and results show improved ADC quantification accuracy in the prostate. This method may be suitable for both clinical imaging and research utilizing ADC quantification.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

Prostate cancer is the leading cause of malignancy in males [1]. Diffusion weighted imaging (DWI) plays an important role in prostate magnetic resonance imaging (MRI) for tumor detection, local-

* Corresponding author at: MR R&D Collaborations, Siemens Healthcare, 10945 Le Conte Ave, Suite 3371J, Los Angeles, CA 90095-7206, USA.
E-mail address: xiaodong.zhong@siemens.com (X. Zhong).

https://doi.org/10.1016/j.jmr.2019.05.007
1090-7807/© 2019 Elsevier Inc. All rights reserved.

ization, and staging [2,3]. However, DWI can suffer from low signal-to-noise ratio (SNR), which degrades image quality, especially the high b-value images. This also impairs quantitative apparent diffusion coefficient (ADC) assessments, where the noise bias is one major confounding factor [4,5].

Typically, high b-value images are measured multiple times and their magnitude-reconstructed images are then averaged to improve the SNR. However, as shown in previous studies [4,5] and later this paper, when the native SNR of each acquired image is inherently low, simply averaging magnitude images does not correct the noise bias, which degrades the accuracy of ADC estimation. If the ADC calculation method does not incorporate the noise distribution, then ADC values calculated using data obtained with only external phased-array coils and high b-values can be underestimated compared to similar measurements obtained using an endorectal coil combined with external phased-array coils [6]. While a number of noise-bias-compensation methods have been described in previous literature, most of these are site-specific techniques as research tools with limited availability. ADC quantification techniques shall be not only accurate and robust, but also broadly commercially available so that they can be reliably used in clinical practice. In addition, most of the literature methods require manual intervention in some steps of the ADC calculation process such as the noise estimation, posing a challenge of utilizing these methods in the clinical environment where a fully automatic calculation is important.

In order to increase SNR, prostate MRI can be performed using an endorectal coil combined with external phased-array coils [6–10]. While adding an endorectal coil provides superior SNR [6–9], it involves extra setup time [10], patient discomfort [11,12], and potential complications such as proctitis and diverticulitis [12,13]. When disposable endorectal coil elements are used, the cost of the disposable components is incurred, while reusable endorectal coils are expensive and must be cleaned thoroughly after use. Another limitation of endorectal coil use is deformation of the prostate, which can be challenging in applications such as radiation therapy planning and fusion biopsy [14]. It has been shown that T2-weighted (T2W) imaging and DWI, two critical components of prostate MRI, can perform similarly for lesion detection and assessment with or without an endorectal coil, suggesting that the use of an endorectal coil may not be always necessary for qualitative diagnosis [11,12]. Recently, it was suggested that using an endorectal coil provided no additional benefit in terms of prostate cancer detection accuracy compared to using external coils only in the context of multiparametric imaging [15]. Ideally, the noise-bias-compensation method should work with external coils without the need to use an additional endorectal coil.

The goal of this study was to develop a fully automatic computational method which runs on a broadly-available hardware platform and is ready to use in the clinical imaging environment for ADC quantification with noise bias compensation, and to perform a preliminary evaluation in prostate DWI. As an extension of a previous study [16], this method was theoretically derived using a direct mapping approach and a stepwise Chebychev polynomial modelling, and implemented as both offline and inline prototypes. A Monte Carlo simulation was performed to assess the theory, followed by a diffusion phantom validation. Since no ground truth was available for direct validation of in vivo data, two studies were performed for preliminary in vivo evaluation. A first in vivo study was performed using acquisitions with external coils only in conjunction with an additional accompanying Monte Carlo simulation to evaluate the consistency between the theoretical and empirical results. In the second vivo study, ADC values obtained with both endorectal and external coils were used as a reference standard to compare to the results of the proposed method obtained with external coils only.

2. Theory

2.1. Diffusion model equation and ADC calculation methods

In DWI, the magnitude signal intensity S_i of the image acquired with the i-th b-value image is modeled as

$$S_i = S_0 e^{-b_i \cdot ADC} \tag{1}$$

where S_0 is the pixel intensity without diffusion weighting. The two unknown variables, S_0 and ADC, can be determined using acquisitions of two or more b-values.

One common method to solve for these unknown variables is log-linear fitting (LL) [17–19]. Denote the acquired signal as $\hat{S}_i$, and given the definitions

$$y_i = \log\left(\hat{S}_i\right) \tag{2}$$

$$x_i = b_i \tag{3}$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \tag{4}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{5}$$

where n is the number of acquired b-value signals. The LL method to calculate ADC in this study was described as

$$\mathrm{ADC} = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \tag{6}$$

Another method is non-linear least squares fitting (LS) [16–18], which seeks to minimize

$$E = \sum_{i=1}^{n} \left|\hat{S}_i - F_i\right|^2 \tag{7}$$

where F_i can be any function value related to the signal model. The common mono-exponential fitting to calculate ADC is equivalent to setting

$$F_i = S_i \tag{8}$$

The Levenberg-Marquardt algorithm [20], for example, can be used to solve for S_0 and ADC.

2.2. Noise distribution and maximum probability

The magnitude MRI data follows a non-central Chi distribution with the probability density function (PDF) [21,22]

$$PDF(m) = \frac{m^L}{\sigma^2} m_0^{1-L} e^{-\frac{m_0^2+m^2}{2\sigma^2}} I_{L-1}\left(\frac{m m_0}{\sigma^2}\right) \tag{9}$$

where m is the random variable corresponding to the acquired magnitude, L is the effective channel number, m_0 is the theoretical noise-free magnitude equivalent to S_i in Eq. (1), σ is the noise standard deviation (SD) of the normal distribution, and I_{L-1} is the $(L-1)$th-order modified Bessel function of the first kind [22]. When L equals 1, Eq. (1) simplifies to a Rician distribution [22,23]. If multi-channel data are combined with the spatial-matched-filter algorithm (or alternatively termed as adaptive coil combination), as used in this study, the magnitude and the real/imaginary part of the channel-combined complex data follow Rician and Gaussian distributions, respectively, i.e. L can be equivalently regarded as 1 [23].

Among several parameters reflecting the PDF characteristics [5], maximum probability (MP) was calculated in this study, which

was identified as the m_{MP} value corresponding to the maximum of the PDF curve

$$PDF(m_{MP}) = \max\{PDF(m)|_{m=0}^{\infty}\} \quad (10)$$

2.3. Noise standard deviation map

While some noise estimation algorithms assume a single noise SD across the image [24], a pixel-by-pixel noise SD map is preferred because factors such as parallel imaging can make such an assumption invalid [23–28]. The noise SD map can be calculated based on analytical equations or models [25–27], or multiple true or pseudo replicas, i.e. repeated acquisitions [23,28]. While analytical methods offer computational efficiency, they depend on specific modeling of individual reconstruction and post-processing procedures. In contrast, independent of the reconstruction and post-processing, replica methods are more flexible. The true replica methods perform repeated acquisitions of the same imaging protocol, which is prone to prolonged acquisitions and motion during the acquisitions [23]. In this study, a pseudo replica method was utilized, which fed the single-acquisition data through the reconstruction multiple times with different noise realizations added and calculated the noise SD map [29,30].

2.4. Noise bias compensation

Iterative algorithms can be used to solve the problem, but are usually computationally-expensive because of the costly calculation of I_{L-1} and the need to compute the noise-biased value based on the theoretical unbiased value in multiple iterations [5,26].

Alternatively, the estimation can be achieved by mapping the noise-biased magnitude value back to the theoretical unbiased magnitude value [27,31], followed by the ADC calculation. Koay et al proposed a simultaneous estimation of the theoretical unbiased magnitude value and noise SD from the noise-biased magnitude value [31], and later an iterative method to estimate them separately [27]. The limitations of the former method include relatively slow convergence at low SNR [31], and the latter method is computationally expensive due to the need for smoothing, iterative estimation and composition of the inverse cumulative probability function [27].

In this work, we propose a direct mapping approach for noise bias compensation. In detail, the magnitude values are first normalized as

$$m'_{MP} = \frac{m_{MP}}{\sigma} \quad (11)$$

A stepwise Chebychev polynomial modelling is approximated to perform a direct and efficient inverse mapping from m'_{MP} to m'_0.

$$m'_{MP} \xrightarrow{\text{stepwise Chebychev polynomial modelling}} m'_0 \quad (12)$$

Lastly, the theoretical unbiased value is scaled back to the original unit.

$$m_0 = m'_0 \cdot \sigma \quad (13)$$

Fitting methods can be subsequently applied to obtain noise-unbiased ADC values.

3. Methods

3.1. Theoretical computer simulation

A Monte Carlo simulation was programmed using in-house code in Matlab (Mathworks, Natick, MA). ADC values were calculated using four different methods, including (a) LL fitting on averaged b-value data (LL AveB); (b) LS fitting on averaged b-value data (LS AveB); (c) LS fitting on non-averaged b-value data (LS Non-AveB); and (d) LS fitting on non-averaged b-value data with MP noise compensation (LS MP-Cor Non-AveB). The LL AveB method is a conventional method, and the LS MP-Cor Non-AveB method is the proposed method. LS AveB and LS Non-AveB were tested as methods involving intermediate changes from LL to LS and from averaging to non-averaging based on the LL AveB method [4,5].

Ground-truth values were set for S0, b-values, a range of b0 image SNR and a range of ADC, and the latter two were two iterative variables for simulation. For each Monte Carlo repetition, the real and imaginary parts of the generated complex signal data were added with random Gaussian noise whose SD was determined by SNR and S0. The magnitude was calculated as the input data for the four different ADC fitting methods. Error percentage (error%) was calculated as

$$Error\% = \frac{Resultant\ \ fitted\ \ ADC - Ground - truth\ \ ADC}{Ground - truth\ \ ADC} \times 100\% \quad (14)$$

With 500 Monte Carlo repetitions, error% curves as a function of SNR were calculated for ADC of 0.8, 1.5 and 3.0×10^{-3} mm^2/s, respectively. The difference between the LS MP-Cor Non-AveB method and the other three methods was evaluated using a one sample *t*-test [32]. Additionally, mean and SD maps of error% were generated as a function of SNR and ADC using the SNR ranging from 0.1 to 10 with a step size of 0.1, and the ADC ranging from 0.1×10^{-3} mm^2/s to 3.5×10^{-3} mm^2/s with a step size of 0.1×10^{-3} mm^2/s, respectively, which cover most practical values that could occur with the clinical protocol. The simulated b-values were 50, 400 and 800 s/mm^2 with repetitions of 2, 4 and 8, respectively. The number of Monte Carlo repetitions was 120 for each SNR-ADC pair.

3.2. Pulse sequence implementation and diffusion phantom validation

A conventional spin-echo (SE) echo-planar imaging (EPI) diffusion-weighted sequence using the LL AveB was prototypically modified to implement the proposed LS MP-Cor Non-AveB method in the image reconstruction framework capable of both inline and offline reconstruction.

A Quantitative Imaging Biomarkers Alliance (QIBA) diffusion phantom (High Precision Devices, Boulder, Colorado) was imaged using 18-channel body and 32-channel spine coils (referred as "external coils" throughout this paper) on a 3T MRI system (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) with parameters in Table 1, including b-values ranging from 0 to 4000 s/mm^2. The phantom is a plastic sphere with a 194-mm diameter, consisting of 13 vials filled with various concentrations

Table 1
Imaging parameters of the diffusion phantom acquisition. Abbreviations: GRAPPA - Controlled aliasing in volumetric parallel imaging [33].

Imaging parameters	Values
TR (ms)	3000
TE (ms)	400
Pixel size (mm^2)	2.4 × 2.4
Slice thickness (mm)	1
Parallel imaging acceleration factor	GRAPPA 2
b-values (s/mm^2)	0, 50, 100, 400, 600, 800, 1200, 1600, 2000, 3000, 4000
Repetitions of the corresponding b-values	16 averages per b-value
Diffusion directions of the corresponding b-values	1 direction for b = 0; 3 directions for all other b-values

of the polymer polyvinylpyrrolidone (PVP) in aqueous solution, which are recommended by the National Institute of Standards and Technology for assessment and standardization of ADC.

The phantom was prepared and maintained in an ice-water bath condition according to the manufacturer's instruction [34], and temperatures were recorded before and after the acquisition. Raw data were saved for offline processing. From the raw data, two b-value sets were extracted and reconstructed for ADC calculation, including one using data from b = 50 and 800 s/mm^2 only, and the other using the full dataset. The mean ADC results of each PVP concentration were averaged using the ADC values of the vials with the same concentrations. The results calculated with both the LL AveB and LS MP-Cor Non-AveB methods were compared to the ground-truth values provided by the manufacturer [34].

3.3. The first in vivo study: validation compared with simulation

This prospective in vivo study was approved by our Institutional Review Board (IRB) and compliant with the Health Insurance Portability and Accountability Act (HIPAA). With written informed consent obtained, 29 males (67.8 ± 7.4 years) undergoing clinically required prostate MRI exams were scanned with external coils only on a 3T MRI scanner (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) different from the scanner used in the phantom experiment, using the same prototypical SE EPI diffusion-weighted sequence in the phantom experiment and protocol 1 in Table 2.

The raw data were saved for offline processing, which allowed calculating the noise SD and SNR maps of the corresponding data in addition to the ADC maps by both the LL AveB and the LS MP-Cor Non-AveB methods. Prostate peripheral and central regions of interest (ROIs) were manually segmented by a MRI physicist with 10 years of experience in abdominal MRI on two adjacent slices of the ADC maps, leading to the data from a total of 116 ROIs to analyze.

An additional Monte Carlo simulation was performed to evaluate the consistency between the theoretical and empirical results. Taking the values from the SNR maps and from the ADC maps measured by the LS MP-Cor Non-AveB method as part of the simulation input parameters and ground-truth values, the simulation was performed with 40 repetitions and generated the simulated ADC results using both the LL AveB and the LS MP-Cor Non-AveB methods.

Table 2
Imaging parameters of the acquisitions of in vivo subjects using different coil configurations and protocols. Abbreviations: GRAPPA - Controlled aliasing in volumetric parallel imaging [33].

Imaging parameters	In Vivo	
	Protocol 1	Protocol 2
Coils	External coils only	Both endorectal and external coils
TR (ms)	5700	5500
TE (ms)	67	68
Pixel size (mm^2)	1.93 × 1.93	1.25 × 1.25
Slice number	32	32
Slice thickness (mm)	3	3
Parallel imaging acceleration factor	GRAPPA 2	GRAPPA 2
b-values (s/mm^2)	50, 400, 800	50, 800
Repetitions of the corresponding b-values	2, 4, 8	2, 4
Diffusion directions of the corresponding b-values	4, 4, 4	3, 3
Acquisition time (s)	336	105

Prior to examining the data, the region of practical equivalence (ROPE) was defined as an absolute value of difference $<0.05 \times 10^{-3}$ mm^2/s, meaning that if two methods differ by $<0.05 \times 10^{-3}$ mm^2/s then the methods are considered practically equivalent. This value was obtained by considering a conservatively strict average difference of about 5% near a value of clinical interest of 1.0×10^{-3} mm^2/s.

Statistical analyses were performed using R v3.3.2 (R Core Team, Vienna, Austria) [32]. The quantiles of the mean values of the measured ADC in all ROIs were compared to the quantiles of those of the simulated ADC in a quantile-quantile (QQ) plot [35]. A two-sample Kolmogorov-Smirnov test was performed to detect the statistical difference between the measured and the simulated ADC distributions [36]. In addition, a linear model was run to test if there was any significant difference between measured and simulated ADC results when accounting for the input SNR, the input ADC, and the fitting method. Only main effects were considered. And $p < 0.05$ was regarded as significantly different.

3.4. The second in vivo study: comparison with different coil configurations

Under the same IRB-approved protocol, 15 additional males undergoing clinical prostate MRI exams were scanned with written informed consent obtained on the same 3T MRI system (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) and the same prototypical sequence as in the first in vivo study. Diffusion-weighted images and ADC maps were reconstructed inline and saved as DICOM images for offline analysis. One subject was excluded because of missing images due to software issues, resulting in a total of 14 males (63.0 ± 4.8 years).

The subjects underwent routine prostate MRI protocols including T2W imaging. As a comparison reference standard, a conventional SE EPI diffusion-weighted sequence was performed with protocol 2 in Table 2, a clinical protocol in our institution utilizing the same external coils and an additional endorectal coil (Medrad eCoil, Bayer, Whippany, NJ). Separately and without the endorectal coil, the modified SE EPI diffusion-weighted sequence was acquired using external coils only with protocol 1 in Table 2 for both the LL AveB and LS MP-Cor Non-AveB methods.

Peripheral and central ROIs were manually segmented by a board-certified radiologist with 11 years of abdominal MRI experience on two adjacent slices of the ADC maps acquired with both endorectal and external coils. Because only one subject had visible lesions, lesion regions were excluded. Additional ROIs were manually segmented and visually co-registered on the ADC maps acquired with only external coils and the LL AveB method, then duplicated onto the ADC maps calculated with the LS MP-Cor Non-AveB method. When necessary, T2W images were used for anatomic reference. Mean and SD of ADC were recorded based on the various ADC maps.

Using R v3.3.2, Bland-Altman was performed [37]. In addition, a linear model was fitted with the dependent variable being the bias and independent variables being the ADC fitting methods and the prostate regions. Only main effects were tested, and $p < 0.05$ was regarded as significantly different with correction for multiple comparisons. Consistently, prior to examining the data, the ROPE was defined as an absolute value of bias $<0.05 \times 10^{-3}$ mm^2/s. To calculate the posterior probability that the true mean bias is within the ROPE [38], a Bayesian regression with an uninformative prior was used to obtain a posterior probability (Bayesian p-value) and a highest posterior density interval (Bayesian confidence interval) [39].

4. Results

4.1. Theoretical computer simulation results

Example error% vs SNR curves for ground-truth ADC = 0.8×10^{-3}, 1.5×10^{-3} and 3.0×10^{-3} mm^2/s are show in Fig. 1, where the error% of the proposed LS MP-Cor Non-AveB method is significantly different from the other methods as evaluated by the one sample *t*-test ($p < 2.2 \times 10^{-16}$).

In Fig. 2A, the mean error% as a function of SNR and ADC shows the largest area with small error values ("Green Zone") for LS MP-Cor Non-AveB. The SD of the error% predicts relatively more noise in the measured ADC results with LS MP-Cor Non-AveB compared to the other methods (Fig. 2B).

4.2. Diffusion phantom results

The diffusion phantom temperature during the acquisition was within the range of 0–0.2 °C. The ADC results from the phantom are shown in Fig. 3. Compared to the ground truth ADC values (black bars in Fig. 3A), when high b-value data were included, LL AveB underestimated ADC values, and the underestimated error % values were greater than their error bars for either relatively large ADC values (above 1.0×10^{-3} mm^2/s, i.e. the normal prostate ADC value

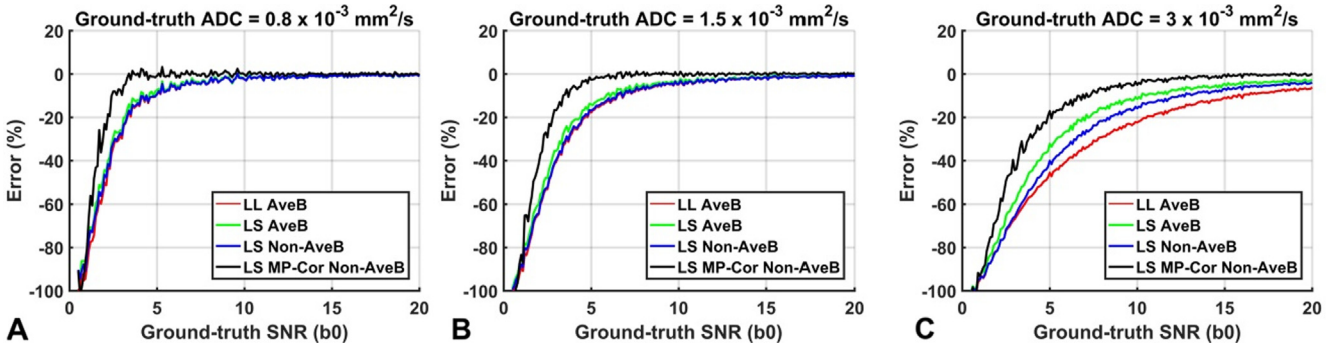

Fig. 1. Error%-SNR curves for three typical ADC values using four different ADC fitting methods. The proposed simulation approach was performed with b = 50, 400, 800 s/mm^2 with 2, 4 and 8 repetitions/averages, respectively. SNR ranged from 0.1 to 20 with a step size of 0.1. The repetition number of Monte Carlo simulation was 500.

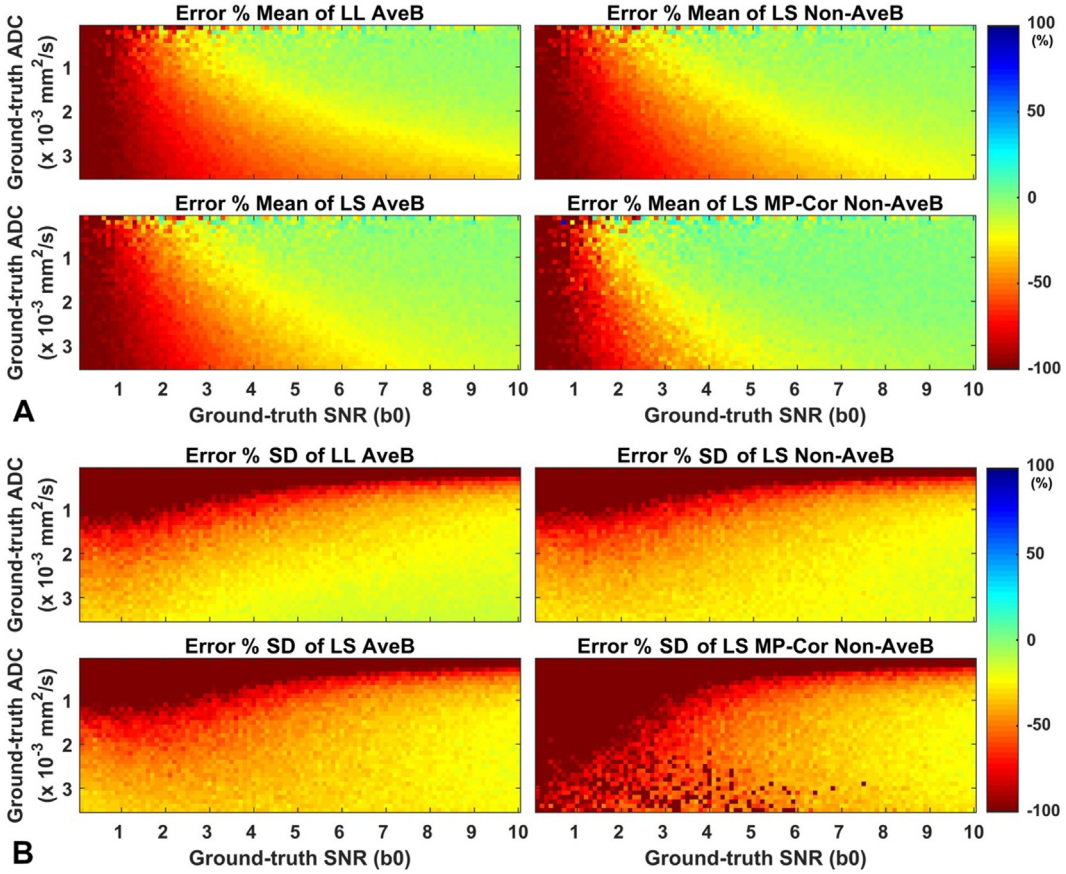

Fig. 2. Mean (A) and SD (B) of error% as a function of ADC and SNR using four different ADC fitting methods. The proposed simulation approach was performed with b = 50, 400, 800 s/mm^2 with 2, 4 and 8 repetitions/averages, respectively. SNR ranged from 0.1 to 10 with a step size of 0.1. ADC ranging from 0.1 to 3.5×10^{-3} mm^2/s with a step size of 0.1×10^{-3} mm^2/s. The repetition number of Monte Carlo simulation was 120.

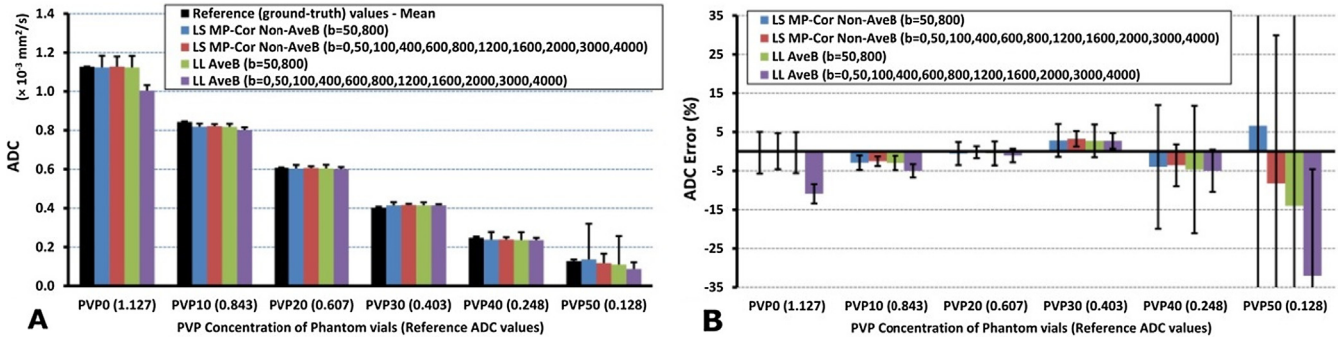

Fig. 3. A: The ADC results of the 5 different PVP concentrations in the diffusion phantom measured with both the LL AveB and the LS MP-Cor Non-AveB for two different b-value sets: [1] b = 50 and 800 s/mm^2, and [2] b = 0, 50, 100, 400, 600, 800, 1200, 1600, 2000, 3000 and 4000 s/mm^2. The ground-truth ADC value of each PVP concentration provided by the phantom manufacturer is listed in the parentheses and also presented as the black bars. The error bars reflect the standard deviation values, which were provided by the phantom manufacturer based on multiple MR spectroscopy measurements for the reference values, and calculated based on different pixels within the measured ROIs for the non-reference imaging methods, respectively. B: The corresponding error percentage plot. The error bars are calculated based on the propagation of the standard deviation values of the non-reference imaging methods and reference values. The results of b-value set [1] exhibit larger error bars (blue and green bars) due to fewer total repetitions compared to the b-value set [2] (red and purple bars). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

range [the purple bar of PVP0 in Fig. 3]) or relatively high concentration of PVP (the purple bar of PVP50 in Fig. 3). The proposed LS MP-Cor Non-AveB method estimated ADC values more consistently with the ground truth values, even in the presence of high b-value (low SNR) data (red bars in Fig. 3).

4.3. The first in vivo study: validation compared with simulation

As shown in Fig. 4A and B, the measured ADC of both the LL AveB and the LS MP-Cor Non-AveB methods provided similar results as the simulated ADC values, demonstrating the consistency between the theoretical and empirical results. A consistent distribution between the measured and simulated ADC results from both the LL AveB and the LS MP-Cor Non-AveB methods also appeared in the QQ plot (Fig. 4C). Additionally, the two-sample Kolmogorov-Smirnov test did not detect any significant difference ($p = 0.5618$) between the measured and the simulated ADC distributions.

As for the linear model, the overall model was significant ($p < 0.0001$) as was each of the effects ($p < 0.0001$) with an adjusted R^2 of 0.9594. The estimate was that the simulated ADC results were on average 0.019×10^{-3} mm^2/s larger than the measured results with a 95% CI of $[0.010, 0.029] \times 10^{-3}$ mm^2/s lying well within the ROPE. Overall, this again indicated good agreement between the simulated and the measured values.

4.4. The second in vivo study: comparison with different coil configurations

Fig. 5 shows one set of example images. The SNR difference between b50 images (Fig. 5A vs. E) with different coil settings is evident, leading to quantitative differences in ADC maps (Fig. 5B, C vs. F). For all the 14 subjects in this study, the in vivo ADC reference values with both endorectal and external coils were $1.45 \pm 0.15 \times 10^{-3}$ mm^2/s in the central zone and $1.78 \pm 0.15 \times 10^{-3}$ mm^2/s in the peripheral zone, and the corresponding results with external coils only are $1.34 \pm 0.16 \times 10^{-3}$ mm^2/s and $1.65 \pm 0.20 \times 10^{-3}$ mm^2/s by LL AveB and $1.49 \pm 0.16 \times 10^{-3}$ mm^2/s and $1.81 \pm 0.20 \times 10^{-3}$ mm^2/s by LS MP-Cor Non-AveB, respectively. One observation is that, the ADC map by the proposed method (Fig. 5C) exhibits relatively larger SD than the conventional method assuming Gaussian noise (Fig. 5B,F). Bland-Altman plots show similar results, with a mean bias of -0.12×10^{-3} mm^2/s for LL AveB and 0.04×10^{-3} mm^2/s for LS MP-Cor Non-AveB (Fig. 6).

The linear model fit showed the ADC fitting algorithms had a highly significant impact on the bias ($p < 0.0001$) while the regions chosen in the prostate did not ($p = 0.818$). LL AveB had significant bias of -0.12×10^{-3} mm^2/s with a 95% confidence interval (CI) of $[-0.17, -0.06] \times 10^{-3}$ mm^2/s ($p < 0.0001$), while LS MP-Cor Non-AveB had non-significant bias of 0.04×10^{-3} mm^2/s with a 95% CI of $[-0.01, 0.09] \times 10^{-3}$ mm^2/s ($p = 0.187$). The bias difference

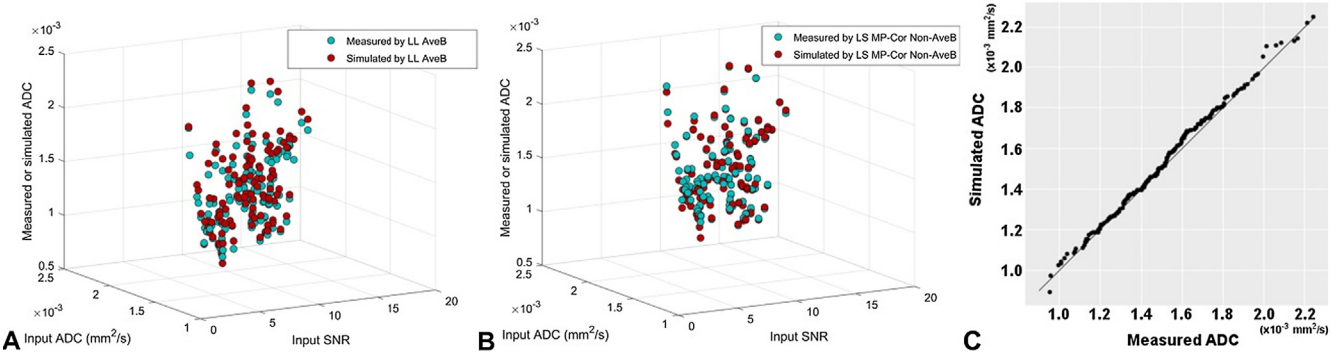

Fig. 4. The comparison of the measured and simulated ADC results of both the LL AveB (A) and the LS MP-Cor Non-AveB methods (B), illustrating prostate peripheral and central regions on two adjacent slices from 29 subjects (116 ROIs totally) of the first in vivo study. The repetition number of Monte Carlo simulation was 40. The QQ plot of the measured and simulated ADC results from both methods is in (C).

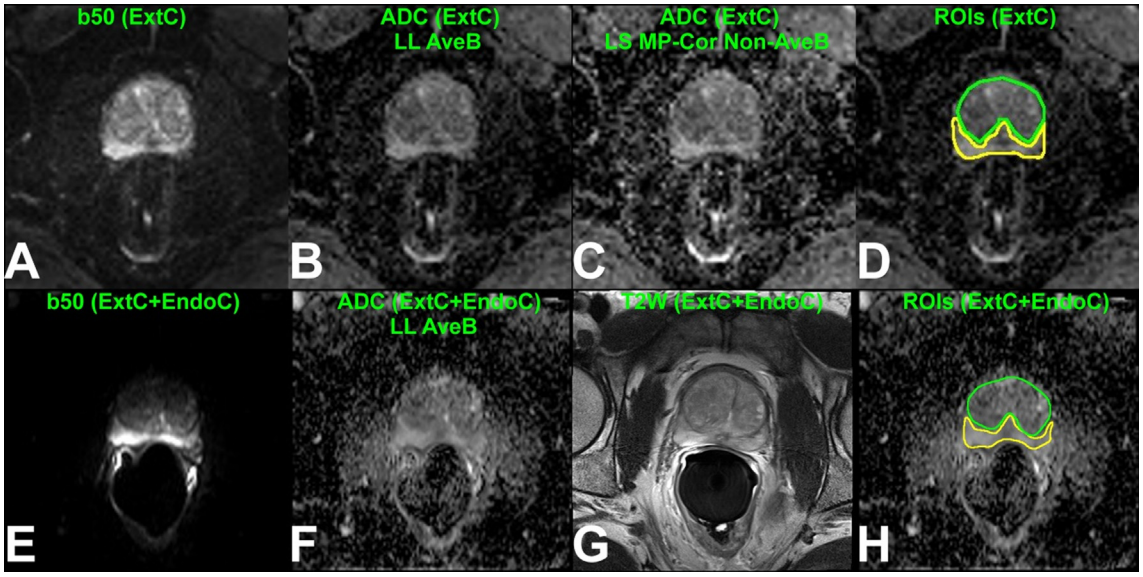

Fig. 5. Representative images and ADC maps obtained using external coils only (A–D) and using the combination of both endorectal and external coils (E–H). The green and yellow ROIs represent the central and peripheral regions of the prostate, respectively (D, H). The images in (A–D) are cropped from the original images for better visual comparison with the images in (E–H). Data were from one patient of the second in vivo study. Abbreviations: ExtC – external coils only; ExtC + EndoC – external and endorectal coils. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

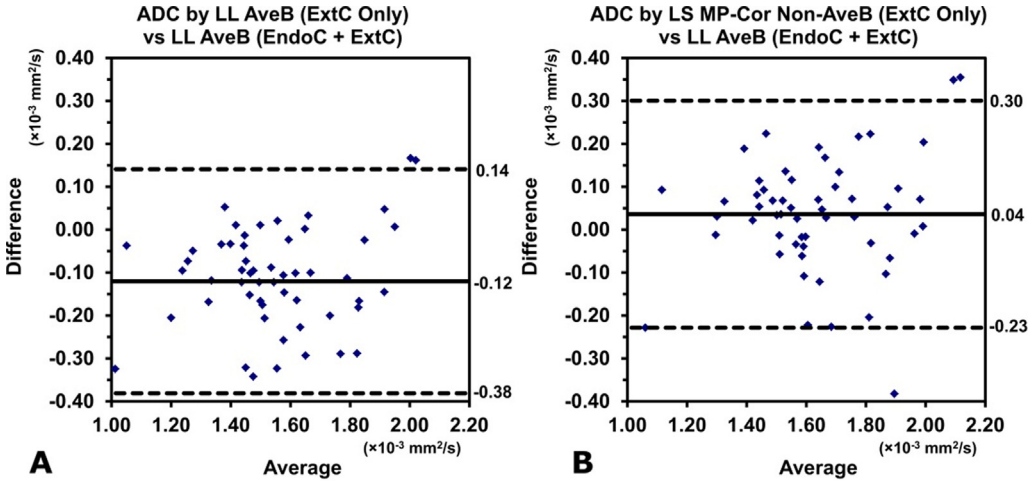

Fig. 6. Bland-Altman plots of in vivo ADC values of prostate peripheral and central regions on two adjacent slices in 14 patients (56 ROIs totally) of the second in vivo study for each of the two different ADC calculation methods using external coils only vs. the LL AveB method using the combination of both external and endorectal coils as a reference. Abbreviations: ExtC – external coils only; ExtC + EndoC – external and endorectal coils.

between LL AveB and LS MP-Cor Non-AveB was significant ($p < 0.001$).

The Bayesian p-value that the mean bias for LL AveB was within the ROPE was $p = 0.0004$ with a 95% Bayesian confidence interval from $[-0.156, -0.085] \times 10^{-3}$ mm^2/s lying entirely outside the ROPE. In contrast, the Bayesian p-value for LS MP-Cor Non-AveB was $p = 0.778$ with a 95% Bayesian confidence interval of $[0.001, 0.072] \times 10^{-3}$ mm^2/s which straddles the boundary of the ROPE. The Bayesian p-value that LS MP-Cor Non-AveB had less absolute bias than LL AveB was $p = 0.9991$. Overall, this indicated LL AveB was biased and LS MP-Cor Non-AveB reduced the amount of bias.

5. Discussion

This study showed the noise induced inaccuracy of ADC quantification, and our proposed fully automatic computational method improved the accuracy of ADC quantification. The results of the first in vivo study (and Monte Carlo simulation) from 29 clinical subjects show consistency of the empirical results compared to the simulation predicted values. The results of the second in vivo study indicate that, using ADC results acquired with a clinical protocol utilizing both endorectal and external coils as a reference standard, ADC quantification using only external coils is compromised by noise bias. Data from 14 clinical subjects demonstrate that these noise bias-induced errors can be reduced by the proposed computational method. The Bayesian regression analysis provides strong evidence of bias in the conventional ADC method as well as moderate evidence that this bias is practically eliminated with the proposed method. Our study shows that it is almost certainly ($p = 0.9991$) reduced with LS MP-Cor Non-AveB compared to LL AveB. This finding merits further prospective assessment with larger number of subjects.

High b-value images often suffer from low SNR. The noise in MRI magnitude data has a Rician distribution and becomes asymmetric at low SNR [21,22]. If not compensated, this asymmetric distribution causes a systematic overestimation of signal intensity at high b-values and consequent underestimation of ADC values [4,5,40]. Previous attempts to improve ADC estimation can be categorized into two main groups: attempts to correct the noise bias by replacing the diffusion model with its approximate expectation under the assumption of Rician noise models [5,25,26]; and transforming the biased diffusion signal to a bias-corrected signal [27,40]. A detailed description of various methods in previous literature is out of scope of this paper. Review of some techniques can be found in the work of Veraart et al. [26], Dikaios et al. [41] and Fusco et al. [15].

Our proposed method falls into the second category, i.e. the mapping method. Our method uses a stepwise Chebychev polynomial modeling equation to directly calculate the mapping results, instead of iterative processes as used in the previously published methods [5,26,27,31]. The polynomial modeling equation has the advantages of being computationally efficient and free of convergence problems. In addition, our proposed method uses a pseudo replica approach [29,30] for accurate pixelwise noise SD map estimation and the spatial-matched-filter algorithm (alternatively termed adaptive coil combination) for multi-channel coil combination, which allows the model to treat multi-channel data in the same way as single-channel data [27,31] and therefore greatly simplifies the estimation problem. In order to validate our proposed method, we used a variety of reference standard ADC values, including the ground-truth values in simulation, the manufacturer-provided ground-truth values of a commercially available diffusion phantom, and the prostate ADC values acquired with both external and endorectal coils in vivo. Comparison of our proposed method to various methods in the previous literature as reference standards is not a focus of this study, and future separate studies are needed to further investigate this topic.

Our method was designed and implemented to be fully automatic, practically useful in both clinical and research settings. The first in vivo study used offline image reconstruction. In contrast, the image reconstruction was completed inline in the clinical environment for the second in vivo study. Although the image reconstruction times were not recorded, they approximately ranged from 1 to 3 min depending on the protocol settings, which is acceptable in clinical practice. The reconstruction times could be further improved by using fast graphics processing units. In addition, using analytical methods [25–27] to calculate the noise SD maps could potentially further reduce the total reconstruction time.

Averaging multiple data acquisitions is a common strategy to increase SNR, especially for high b-value data. However, this increases data acquisition time substantially. Moreover, as showed in this study and in previous literature [5], when each b-value image's SNR is inherently low, simply averaging magnitude data does not correct the noise bias and does not improve the accuracy of ADC estimation. Only if the SNR of each b-value image is sufficient will the conventional methods, assuming a symmetric (Gaussian) noise distribution, yield accurate ADC quantification. To increase the SNR of high b-value images, one may use dedicated hardware such as endorectal coils, or pulse sequences which inherently provide higher SNR such as segmented EPI [42]. The proposed method compensates for the asymmetric noise distribution without introducing the workflow issues associated with using an endorectal coil or the time penalty of a segmented EPI sequence. Such a method also avoids the issues of patient discomfort and gland distortion associated with endorectal coil placement. Alternatively, complex signal averaging may provide another method for retaining the Gaussian noise distribution for ADC quantification, assuming that reliable phase correction algorithms are applied to correct the motion-induced phase shifts [5]. Another possibility is to work with real and imaginary data to suppress the noise bias as shown in a recent study using principal component analysis [43].

Although designed and validated in the context of prostate DWI, this method could theoretically be extended to applications which use magnitude data as inputs and suffer from the same noise bias issue. Possible applications include ADC mapping in other organs and tissues, proton density fat fraction mapping, T1 mapping, T2/T2* mapping and other quantitative imaging techniques.

Since the purpose of this study was to report our proposed method and perform a preliminary evaluation, some topics remain to be examined and investigated. For example, the ADC values tested in vivo fell mainly within a range of $1.0–2.0 \times 10^{-3}$ mm^2/s. While this covers the ADC values for the normal prostate peripheral and central regions, high grade prostate cancer ADC values may fall below 1.0×10^{-3} mm^2/s. Another topic is that, although it corrected for the effect of an asymmetric noise distribution for a large range of ground-truth ADC and SNR, the ADC results exhibit increased SD especially for the surrounding tissues (Fig. 5C). This observation is predicted in the computer simulation (Fig. 2B), and is consistent with previous results (Fig. 12 in Ref. [25]). It should be noted that evaluating ADC within an ROI helps reduce the effect of the noise in the ADC maps calculated using the proposed method [25]. Lastly, our proposed method worked well within the range of b-values tested in this study, i.e. 0, 50, 100, 400, 600, 800, 1200, 1600, 2000, 3000 and 4000 s/mm^2 in a diffusion phantom, and 50, 400 and 800 s/mm^2 in clinical prostate DWI. Its performance with very high b values to a level that DWI signals are overwhelmed by noise remains to be evaluated in future studies.

6. Conclusion

In this study, a fully automatic computational method to compensate for noise-related bias is proposed to improve the accuracy of ADC quantification, and was validated in simulation and a diffusion phantom. Our first in vivo prostate DWI study shows the consistency between the theoretical and empirical results using a clinically routine protocol and an accompanying simulation. The findings of our second in vivo prostate DWI study suggest that, compared to the reference standard ADC results measured with a clinically relevant protocol utilizing both external coils and an endorectal coil, this method provides improved accuracy of ADC results than the conventional method when only external coils were used. This method may improve quantitative prostate diffusion imaging when noise bias exists. While further validation is needed, this method may be suitable for both clinical imaging and research utilizing ADC quantification.

Acknowledgement

The work performed in this study was supported in part by a grant from Siemens Healthcare (M.B.). The authors gratefully acknowledge Alto Stemmer, PhD for the contribution of the pulse sequence, Vibhas Deshpande, PhD for handling of the diffusion phantom, Elisabeth Weiland, PhD for discussion of the prostate DWI protocols, and Peng Hu, PhD for discussion of manuscript preparation.

Conflict of interests

X.Z., B.D., M.N., S.K. and B.K. are Siemens employees. Institutions have two pending patents describing the proposed method in this study.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics 2016, CA Cancer J. Clin. 66 (2016) 7–30.

[2] H.A. Vargas, O. Akin, T. Franiel, Y. Mazaheri, J. Zheng, C. Moskowitz, K. Udo, J. Eastham, H. Hricak, Diffusion-weighted endorectal MR imaging at 3 T for prostate cancer: tumor detection and assessment of aggressiveness, Radiology 259 (2011) 775–784.

[3] T. Kobus, P.C. Vos, T. Hambrock, M. De Rooij, C.A. Hulsbergen-Van de, J.O. Kaa Barentsz, A. Heerschap, T.W. Scheenen, Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T, Radiology 265 (2012) 457–467.

[4] D.K. Jones, P.J. Basser, Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data, Magn. Reson. Med. 52 (2004) 979–993.

[5] A. Kristoffersen, Optimal estimation of the diffusion coefficient from non-averaged and averaged noisy magnitude data, J. Magn. Reson. 187 (2007) 293–305.

[6] Y. Mazaheri, H.A. Vargas, G. Nyman, A. Shukla-Dave, O. Akin, Hricak H, Diffusion-weighted MRI of the prostate at 3.0 T: comparison of endorectal coil (ERC) MRI and phased-array coil (PAC) MRI-The impact of SNR on ADC measurement, Eur. J. Radiol. 82 (2013) e515–e520.

[7] K.K. Yu, H. Hricak, R. Alagappan, D.M. Chernoff, P. Bacchetti, C.J. Zaloudek, Detection of extracapsular extension of prostate carcinoma with endorectal and phased-array coil MR imaging: multivariate feature analysis, Radiology 202 (1997) 697–702.

[8] M.D. Schnall, Y. Imai, J. Romaszewski, H.M. Pollack, R.E. Lenkinski, H.Y. Kressel, Prostate cancer: local staging with endorectal surface coil MR imaging, Radiology 178 (1991) 797–802.

[9] A.V. D'Amico, R. Whittington, M. Schnall, et al., The impact of the inclusion of endorectal coil magnetic resonance imaging in a multivariate analysis to predict clinically unsuspected extraprostatic cancer, Cancer 75 (1995) 2368–2372.

[10] S.M. Noworolski, J.C. Crane, D.B. Vigneron, J. Kurhanewicz, A clinical comparison of rigid and inflatable endorectal-coil probes for MRI and 3D MR spectroscopic imaging (MRSI) of the prostate, J. Magn. Reson. Imag. 27 (2008) 1077–1082.

[11] A.D. Baur, T. Daqqaq, M. Wagner, A. Maxeiner, A. Huppertz, D. Renz, B. Hamm, T. Fischer, T. Durmus, T2- and diffusion-weighted magnetic resonance imaging at 3T for the detection of prostate cancer with and without endorectal coil: an intraindividual comparison of image quality and diagnostic performance, Eur. J. Radiol. 85 (2016) 1075–1084.

[12] S.H. Lee, K.K. Park, K.H. Choi, B.J. Lim, J.H. Kim, S.W. Lee, B.H. Chung, Is endorectal coil necessary for the staging of clinically localized prostate cancer? Comparison of non-endorectal versus endorectal MR imaging, World J. Urol. 28 (2010) 667–672.

[13] B. Turkbey, P.S. Albert, K. Kurdziel, P.L. Choyke, Imaging localized prostate cancer: current approaches and new developments, AJR Am. J. Roentgenol. 192 (2009) 1471–1480.

[14] J.M. Hensel, C. Menard, P.W. Chung, M.F. Milosevic, A. Kirilova, J.L. Moseley, M. A. Haider, K.K. Brock, Development of multiorgan finite element-based prostate deformation model enabling registration of endorectal coil magnetic resonance imaging for radiotherapy planning, Int. J. Radiat. Oncol. Biol. Phys. 68 (2007) 1522–1528.

[15] R. Fusco, M. Sansone, V. Granata, S.V. Setola, A. Petrillo, A systematic review on multiparametric MR imaging in prostate cancer detection, Infect. Agent. Cancer 12 (2017) 57.

[16] X. Zhong, M.D. Nickel, S.A.R. Kannengiesser, A. Stemmer, B.M. Dale, B. Kiefer, M.R. Bashir, Improved accuracy of apparent diffusion coefficient (ADC) quantification: evaluation in prostate diffusion imaging without using endorectal coils, Proc. Intl. Soc. Mag. Reson. Med. 25 (2017).

[17] J.H. Burdette, A.D. Elster, P.E. Ricci, Calculation of apparent diffusion coefficients (ADCs) in brain using two-point and six-point methods, J. Comput. Assist. Tomogr. 22 (1998) 792–794.

[18] N. Yoshino, I. Yamada, N. Ohbayashi, E. Honda, M. Ida, T. Kurabayashi, K. Maruyama, T. Sasaki, Salivary glands and lesions: evaluation of apparent diffusion coefficients with split-echo diffusion-weighted MR imaging - initial results, Radiology 221 (2001) 837–842.

[19] H.C. Thoeny, F. De Keyzer, C. Boesch, R. Hermans, Diffusion-weighted imaging of the parotid gland: Influence of the choice of b-values on the apparent diffusion coefficient value, J. Magn. Reson. Imag. 20 (2004) 786–790.

[20] K. Levenberg, A method for the solution of certain non-linear problems in least squares, Q. Appl. Math. 2 (1944) 164–168.

[21] H. Gudbjartsson, S. Patz, The Rician distribution of noisy MRI data, Magn. Reson. Med. 34 (1995) 910–914.

[22] C.D. Constantinides, E. Atalar, E.R. McVeigh, Signal-to-noise measurements in magnitude images from NMR phased arrays, Magn. Reson. Med. 38 (1997) 852–857, Erratum Magn. Reson. Med. 52 (2004) 219.

[23] O. Dietrich, J.G. Raya, S.B. Reeder, M. Ingrisch, M.F. Reiser, S.O. Schoenberg, Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics, Magn. Reson. Imag. 26 (2008) 754–762.

[24] R.M. Henkelman, Measurement of signal intensities in the presence of noise in MR images, Med. Phys. 12 (1985) 232–233.

[25] J.L. Andersson, Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: why, how and but, Neuroimage 42 (2008) 1340–1356.

[26] J. Veraart, J. Rajan, R.R. Peeters, A. Leemans, S. Sunaert, J. Sijbers, Comprehensive framework for accurate diffusion MRI parameter estimation, Magn. Reson. Med. 70 (2013) 972–984.

[27] C.G. Koay, E. Ozarslan, P.J. Basser, A signal transformational framework for breaking the noise floor and its applications in MRI, J. Magn. Reson. 197 (2009) 108–119.

[28] M.J. Riffe, M. Blaimer, K.J. Barkauskas, J.L. Duerk, M.A. Griswold, SNR estimation in fast dynamic imaging using bootstrapped statistics, Proc. Intl. Soc. Mag. Reson. Med. 15 (2007).

[29] P.M. Robson, A.K. Grant, A.J. Madhuranthakam, R. Lattanzi, D.K. Sodickson, C.A. McKenzie, Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions, Magn. Reson. Med. 60 (2008) 895–907.

[30] C.N. Wiens, S.J. Kisch, J.D. Willig-Onwuachi, C.A. McKenzie, Computationally rapid method of estimating signal-to-noise ratio for phased array image reconstructions, Magn. Reson. Med. 66 (2011) 1192–1197.

[31] C.G. Koay, P.J. Basser, Analytically exact correction scheme for signal extraction from noisy magnitude MR signals, J. Magn. Reson. 179 (2006) 317–322.

[32] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2016. <https://www.R-project.org/>.

[33] F.A. Breuer, M. Blaimer, M.F. Mueller, N. Seiberlich, R.M. Heidemann, M.A. Griswold, P.M. Jakob, Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA), Magn. Reson. Med. 55 (2006) 549–556.

[34] K.E. Keenan, S. Carnicka, S.C. Gottlieb, M.A. Boss, K.F. Stupic, Assessing changes in MRI measurands incurred in a scanner upgrade: is my study compromised?, Proc Intl. Soc. Mag. Reson. Med. 25 (2017).

[35] M.B. Wilk, R. Gnanadesikan, Probability plotting methods for the analysis of data, Biometrika 55 (1968) 1–17.

[36] M.A. Stephens, EDF statistics for goodness of fit and some comparisons, J. Am. Stat. Assoc. 69 (1974) 730–737.

[37] Bernhard Lehnert, BlandAltmanLeh: Plots (Slightly Extended) Bland-Altman Plots, R package version 0.3.1, 2015. https://CRAN.R-project.org/package=BlandAltmanLeh.

[38] J.K. Kruschke, Bayesian estimation supersedes the t-test, J. Exp. Psychol. Gen. 142 (2013) 573–603.

[39] A.D. Martin, K.M. Quinn, J.H. Park, MCMCpack: Markov Chain Monte Carlo in R, J. Stat. Softw. 42 (2011) 1–21. http://www.jstatsoft.org/v42/i09/.

[40] C.G. Koay, L.C. Chang, J.D. Carew, C. Pierpaoli, P.J. Basser, A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging, J. Magn. Reson. 182 (2006) 115–125.

[41] N. Dikaios, S. Punwani, V. Hamy, P. Purpura, S. Rice, M. Forster, R. Mendes, S. Taylor, D. Atkinson, Noise estimation from averaged diffusion weighted images: can unbiased quantitative decay parameters assist cancer evaluation?, Magn Reson. Med. 71 (2014) 2105–2117.

[42] D.A. Porter, R.M. Heidemann, High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition, Magn. Reson. Med. 62 (2009) 468–475.

[43] N.K. Chen, H.C. Chang, A. Bilgin, A. Bernstein, T.P. Trouard, A diffusion-matched principal component analysis (DM-PCA) based two-channel denoising procedure for high-resolution diffusion-weighted MRI, PLoS One 13 (2018) e0195952.