



Test-retest repeatability of ADC in prostate using the multi b -Value VERDICT acquisition

Harriet J. Rogers^{a,*,1}, Saurabh Singh^a, Anna Barnes^a, Nancy A. Obuchowski^b, Daniel J. Margolis^c, Dariya I. Malyarenko^d, Thomas L. Chenevert^d, Amita Shukla-Dave^e, Michael A. Boss^f, Shonit Punwani^a

^a Centre for Medical Imaging, Division of Medicine, University College London, London, UK

^b Department of Quantitative Health Sciences, Cleveland Clinic Foundation, Cleveland, OH, USA

^c Radiology, Weill Cornell Medicine, New York, NY, USA

^d Department of Radiology, University of Michigan, Ann Arbor, MI, USA

^e Departments of Medical Physics and Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

^f Center for Research and Innovation, American College of Radiology, Philadelphia, PA, USA

ARTICLE INFO

Keywords:

Magnetic Resonance Imaging
Prostate Cancer
Diffusion Weighted Imaging
Apparent Diffusion Coefficient
Repeatability

ABSTRACT

Purpose: VERDICT (Vascular, Extracellular, Restricted Diffusion for Cytometry in Tumours) MRI is a multi b -value, variable diffusion time DWI sequence that allows generation of ADC maps from different b -value and diffusion time combinations. The aim was to assess precision of prostate ADC measurements from varying b -value combinations using VERDICT and determine which protocol provides the most repeatable ADC.

Materials and Methods: Forty-one men (median age: 67.7 years) from a prior prospective VERDICT study (April 2016–October 2017) were analysed retrospectively. Men who were suspected of prostate cancer and scanned twice using VERDICT were included. ADC maps were formed using 5 b -value combinations and the within-subject standard deviations (wSD) were calculated per ADC map. Three anatomical locations were analysed per subject: normal TZ (transition zone), normal PZ (peripheral zone), and index lesions. Repeated measures ANOVAs showed which b -value range had the lowest wSD, Spearman correlation and generalized linear model regression analysis determined whether wSD was related to ADC magnitude and ROI size.

Results: The mean lesion ADC for b_0 b_{1500} had the lowest wSD in most zones (0.18–0.58 $\times 10^{-4}$ mm²/s). The wSD was unaffected by ADC magnitude (Lesion: $p = 0.064$, TZ: $p = 0.368$, PZ: $p = 0.072$) and lesion Likert score ($p = 0.95$). wSD showed a decrease with ROI size pooled over zones ($p = 0.019$, adjusted regression coefficient = -1.6×10^{-3} , larger ROIs for TZ versus PZ versus lesions). ADC maps formed with a maximum b -value of 500 s/mm² had the largest wSDs (1.90–10.24 $\times 10^{-4}$ mm²/s).

Conclusion: ADC maps generated from b_0 b_{1500} have better repeatability in normal TZ, normal PZ, and index lesions.

1. Introduction

Multiparametric MRI (mpMRI) is an established imaging technique for investigating suspected prostate cancer [1]. Diffusion Weighted Imaging (DWI) and the derived Apparent Diffusion Coefficient (ADC)

maps are a central part of the mpMRI protocol. PI-RADS V2.1 (Prostate Imaging-Reporting and Data System) acknowledge that ADC is often calculated from $\geq 2b$ -values and that ADC is affected by b -value choice [2]. If only 2 b -values are acquired, the recommendation is to use 100 s/mm² (preferably 50–100 s/mm²) for the lowest and 800–1000 s/mm²

Abbreviations: ADC, Apparent Diffusion Coefficient; VERDICT, Vascular Extracellular and Restricted Diffusion for Cytometry in Tumors; PI-RADS, Prostate Imaging Reporting and Data System; ROI, Region of Interest; mpMRI, Multiparametric MRI; wSD, Within-subject Standard Deviation; TZ, Transition Zone; PZ, Peripheral Zone; DWI, Diffusion Weighted Imaging.

* Corresponding author at: Guy's & St Thomas' Trust, Westminster Bridge Road, London SE1 7EH, UK.

E-mail addresses: harrietr99@gmail.com, harriet.rogers@gstt.nhs.uk (H.J. Rogers).

¹ Work originated from: Centre for Medical Imaging, Division of Medicine, University College London, Charles Bell House, 43-45 Foley Street, London, W1W 7TS, UK.

<https://doi.org/10.1016/j.ejrad.2023.110782>

Received 19 January 2023; Received in revised form 24 February 2023; Accepted 10 March 2023

Available online 16 March 2023

0720-048X/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

for b_{max} . This range is achievable by most clinical scanners and should avoid the diffusion kurtosis effect, seen at higher b -values. Due to a lack of ADC repeatability and reproducibility data, current PI-RADS uses ADC only as a qualitative component of prostate cancer risk stratification. Recently implemented guidelines by European Association of Urology [3] (EAU, 2019) advocate the use of quantitative ADC thresholds for detection of low-grade cancer to allow $\sim 30\%$ reduction in unnecessary biopsies [4].

For quantitative prostate ADC use, QIBA (Quantitative Imaging Biomarkers Alliance) suggests that $2b$ -values are included: lower b -values: 50–100 s/mm² and $b_{max} = 500$ –1500 s/mm² [5]. QIBA stresses the importance of assessing the precision (repeatability) of ADC to define confidence intervals (CIs) for quantitative ADC thresholds, which can be determined through calculations of the wSD (within-subject standard deviation) of test–retest scans [6]. For longitudinal studies, e.g. assessing therapy response, wSD can be used to determine the significant changes in ADC [5,6]. To establish CI for cross-sectional, diagnostic thresholds [5,6] estimate of bias (accuracy) would be required in addition to repeatability. The range and combination of b -values have an impact on prostate ADC, with significant differences observed from ADC from $2b$ to values compared to multiple b -values [7,8]. Furthermore, b_{max} may impact clinical utility; one study showed greater accuracy at predicting prostate cancer with ADC from $b_{max} = b_{1000}$ compared to $b_{max} = b_{2000}$, whereas the opposite has also been demonstrated [9]. Since ADC values are varying among the prostate zones and lesions (e.g., lower for TZ and cancer), different b -value combinations are expected to have varying impact on ADC contrast and repeatability in different prostate locations.

To date, a handful of rigorous prostate ADC repeatability studies, with limited subject numbers (<12) [10–12], have shown low prostate ADC repeatability ($\sim 50\%$) compared to ADC of other organs [5]. QIBA recommends > 35 subjects for test–retest studies to achieve nominal confidence intervals for repeatability estimate [13]. The larger multicentre ADC repeatability study of 29 subjects in the imaging arm of ACRIN 6701 trial [14], has recently reported encouraging observations of better than 10% ADC ($b = 0,800$ s/mm²) repeatability for the whole prostate scanned on the same day or up to two weeks apart, but included evaluation for only 10 lesions with much lower reported ADC repeatability (up to 40%). Due to small subject numbers and challenges of DWI prostate acquisition and analysis, previous studies were not statistically powered to assess the dependence of ADC repeatability on acquisition protocol (b -values and diffusion times) and analysis (lesion size and location). Such dependencies are of interest for radiologists to improve prostate ADC precision and enable the inclusion of quantitative ADC information in future PI-RADS guidelines.

VERDICT (Vascular, Extracellular, Restricted Diffusion for Cytometry in Tumours) is an advanced diffusion imaging technique that utilises six b -values (0 90 500 1500 2000 3000 s/mm²) with varying diffusion times, and a mathematical model to derive quantitative intra- and extra-cellular tissue fractions [15]. VERDICT derived parameters have demonstrated clinical promise to distinguish between high and low grade prostate cancer [16], and to avoid unnecessary biopsies [17]. The acquired diffusion data also allows generation of ADC maps using different permutations of b -values and diffusion times. Given the potential clinical utility, and that the sequence consists of multiple b -values, it would be beneficial to establish if a repeatable ADC could also be derived from the VERDICT scan, as opposed to acquiring an additional DWI sequence separately.

Detection of clinically significant changes in ADC requires determination of the baseline variability in ADC measurements [6]. This study sought to establish which b -value combination provides the most repeatable prostate ADC, using the VERDICT acquisition for test–retest assessment of > 35 subjects.

2. Materials and Methods

The present study retrospectively analysed data from a repeatability arm ($n = 41$) of a prospective cohort study ($n = 70$) called INNOVATE (ClinicalTrials.gov identifier NCT02689271) which focused on the diagnostic utility of VERDICT parameters, rather than ADC repeatability due to varying diffusion times for different b -value combinations [16]. Ethical approval was obtained from the London–Surrey Borders Research Ethics Committee and written informed consent was attained from study participants. The full study protocol is available at [18]. Prostate Cancer UK funded the trial.

2.1. Study participants

Men were eligible if they had clinical suspicion of prostate cancer due to either a raised PSA or a suspicious digital rectal examination, or were undergoing active surveillance. Participants were excluded if they had undergone prostate cancer treatment, were receiving ongoing hormonal prostate cancer treatment, or had received a biopsy within 6 months before their MRI.

The repeatability cohort included 41 men who had two identical multi b -value VERDICT acquisitions with an interval ≤ 5 min (median age: 67.7 years; range: 50–82 years). Of these 41, 10 men were randomly selected to vacate the scanner after their first scan which necessitated repositioning for the second acquisition. The remaining 31 stayed in the scanner between acquisitions and were not repositioned. Patient demographics are listed in Table 1. Analysis was separated for the 2 repositioning groups. All men underwent single clinical mpMRI in addition to the repeated multi b -value VERDICT acquisition on a 3 T MRI scanner. The mpMRI was assessed by experienced uro-radiologists (>10 years' experience in mpMRI).

Following mpMRI, 19 men with a PI-RADS score ≥ 3 had targeted transperineal biopsy of identified lesions. Histologic examinations from the biopsy cores were evaluated in the standard clinical fashion and assigned an overall Gleason Grade [19] (A.F. and M.R, 13- and 15-years prostate pathology experience, respectively).

2.2. MR Imaging and image analysis

All imaging was performed on a 3 T Philips Achieva scanner. Hyoscine butylbromide (Buscopan, Boehringer Ingelheim, Ingelheim am Rhein, Germany; 0.2 mg/kg, up to 20 mg) was intravenously administered prior to imaging to reduce peristalsis.

Details of the multi b -value VERDICT sequence are in Table 2 (Acquisition Time: 12 min 25 s). Patients also underwent mpMRI in the same session [20].

ADC maps were generated from subsets of the multi b -value VERDICT data. Trace images were generated for each b -value. These were registered to b_0 from b_{3000} from the first acquisition for each subject. ADC maps were generated using an in-house model designed in MATLAB (version 2020a, MathWorks Inc., Natick, MA, USA) on a voxel-by-voxel basis using Eq. (1), where $S(b)$ is the signal at a given b -value, and S_0 is the signal with no diffusion weighting.

Table 1
Patient demographics.

	Non-Repositioned Cohort (n = 31)	Repositioned Cohort (n = 10)
Median Age (years)	67	68
Age Range (years)	50 – 79	50 – 81
Median Serum PSA (ng/ml)	8.96	6.25
Serum PSA Range (ng/ml)	2.25 – 30.93	2.50 – 19.83

Note. – PSA = Prostate-Specific Antigen.

Table 2
The VERDICT MRI acquisition.

G mT/m	Δ (ms)	δ (ms)	b value s/mm ²	TE (ms)	TR (ms)	NEX	Acquisition Time (seconds)
61.2	23.8	3.9	90	50	2482	4	42
44.4	31.3	11.4	500	65	2482	6	107
32.1	43.8	23.9	1500	90	2398	6	146
80.5	32.3	12.4	2000	67	3897	6	237
60.2	38.8	18.9	3000	80	3349	6	204

Note. – G = gradient strength, Δ = timing between gradients, δ = gradient pulse duration, TE = echo time, TR = repetition time, NEX = number of excitations.

$$S(b) = S_0 e^{-b \cdot ADC} \quad (1)$$

Various b -value combinations were used to generate ADC maps to test ADC repeatability. In total, 5 versions of ADC maps were created for each subject for scan1 and scan2 (Table 3). These 5 b -value combinations were selected from all available to ensure there was ≤ 15 ms difference between the scan TEs (to alleviate inconsistent T₂ weighting for different b -values intrinsic to the VERDICT protocol, Table 2) and excluded b_{3000} which may not be clinically achievable. TE = 80 ms was used for all combinations which included b_0 , (the b_0 image from the $b_{max} = 3000$ acquisition from the full VERDICT sequence).

Regions of Interest (ROIs) were created using Mango Software (Research Imaging Institute, UTHSCSA). ROIs were drawn in 3 locations per subject: around the index lesion identified on mpMRI, in normal TZ and normal PZ, by an experienced board-certified radiologist (S.S 4 years' experience in mpMRI). As all trace images were registered to the b_0 from b_{3000} from the first acquisition per subject, only one set of ROIs was needed per subject. These ROIs were then applied onto the registered ADC maps for scan2. Normal PZ and TZ were selected based on PI-RADS scores of 1 or 2. Analyses were confined within the ROIs; therefore, if ROI delineation was not possible within a certain location it was excluded for that subject. Furthermore, image quality was visually assessed and if an ROI showed poor quality or artefacts it was also removed from analysis.

2.3. Statistical analysis

Statistical analysis was conducted in SPSS (version 24 [IBM, Armonk, NY]) and SAS, and hypothesis tests were two-sided. The assumption of normality for wSD of the ADC was checked by performing a Wilks Shapiro test, and also by examining the mean, median, and mode of the distribution, as well as skewness and kurtosis. To assess repeatability, the within-subject standard deviation (wSD) [5] was calculated for each subject and each ADC map in all 3 locations (normal TZ, normal PZ, and index lesions). This was achieved by calculating the variance for each subject, given by the squared difference between both measurements divided by 2 and taking the mean of the variances across all subjects (providing the within-subject variance), then taking the square root of this value.

The wSD was calculated for all five ADC maps per region. The distributions of wSD, wSD², and the log distributions were compared and the robustness of the wSD results was assessed with a model based on the logs. To test if repeatability was affected by the magnitude of the ADC

Table 3
The 4 b -value combinations used to form the ADC maps.

ID of Range	b -value Combination (s/mm ²)
1	0 500
2	0 1500
3	0 2000
4	90 500
5	0 500 2000

Note. – ADC = Apparent Diffusion Coefficient.

value, for each b -value combination, GLMs (Generalized Linear Models) were fit where the dependent variable was the wSD and the independent variable was the magnitude of the ADC (using ADC mean from scan1 and scan2). To determine which b -value range provided the lowest wSD, repeated measure ANOVAs with Bonferroni adjustments for multiple comparisons were conducted for the separate locations. GLMs were performed for each region (TZ, PZ and index lesions). First, the interaction between b -value combination and whether or not the patient was repositioned was tested. If statistically significant ($p < 0.05$), these subjects were analysed separately, if not significant, the data was pooled and adjusted for the binary variable of whether or not the patient had been repositioned. A statistically significant difference between wSD values was then considered if $p < 0.05$.

Bland-Altman plots were used to identify trends in the differences and construct limits of agreement (LOA) between the two repeated scans. Additionally, Spearman correlation coefficients were used to assess the association between ADC mean and Lesion Likert score. The associations based on the correlation coefficients were interpreted as: 0.00–0.20 = negligible, 0.21–0.40 = weak, 0.41–0.60 = moderate, 0.61–0.8 = strong, and 0.81–1.00 = very strong. A statistically significant relationship was considered if $p < 0.05$.

To assess the effect of ROI size on repeatability, the data from normal TZ, normal PZ and index lesions were pooled. A GLM was fit where the dependent variable was the log of wSD and the independent variables were location (normal TZ, normal PZ, or index lesion) and voxel count within ROIs. We used GEEs (Generalized Estimating Equations) to account for the clustered nature of the data, treating patients as a random effect and with an exchangeable working correlation matrix structure, and assessed adjusted regression coefficient for ROI voxel count. A Wald test was used to determine if the regression coefficient for number of voxels was significantly different from zero. A statistically significant difference was considered if $p < 0.05$.

3. Results

A total of 41 men were included in the short-term repeatability study (31 who did not undergo repositioning between scan1 and scan2, and 10 who were randomly assigned to be repositioned). A flow chart detailing the patient numbers is shown in Fig. 1. In total, ADC repeatability was studied for 37 prostate regions in normal TZ, 30 in normal PZ and 35 index lesions (19 biopsied).

Of the non-repositioned group, ROI delineation (e.g., Fig. 2) was possible in the normal TZ for 30 participants and in the normal PZ for 26 participants. Twenty-six men had index lesions which were scored using PI-RADS (PI-RADS Version 2.1). All index lesions scored PI-RADS ≥ 3 (14/26 (53.8%) = 3, 9/26 (34.6%) = 4, 3/26 (11.5%) = 5). Four index lesions were defined in the TZ (2/4 were in the non-repositioned cohort and 1 of these overlapped between the PZ and TZ). Sixteen of the 26 lesions underwent targeted biopsy: 6 had no significant cancer, and the remaining 10 had clinically significant prostate cancer with a Gleason Grade $\geq 3 + 4$ (8/10 (80%) = 3 + 4, 1/10 (10%) = 4 + 3, 1/10 (10%) = 4 + 5). The PI-RADS scores for the biopsied lesions were as follows: 7/16 (43.8%) = PI-RADS 3, 6/16 (37.5%) = PI-RADS 4, and 3/16 (18.8%) = PI-RADS 5.

Of the 10 men who were repositioned between scans, ROI delineation was possible in the normal TZ for 7 participants and in the normal PZ for 4 participants. Nine men had index lesions which all had PI-RADS ≥ 3 (3/9 (33.3%) = 3, 4/9 (44.4%) = 4, 2/9 (22.2%) = 5). Three of the 9 lesions underwent targeted biopsy all with clinically significant prostate cancer with Gleason Grade $\geq 3 + 4$ (1/3 (33.3%) = 3 + 4, 1/3 (33.3%) = 4 + 3, 1/3 (33.3%) = 4 + 5). The PI-RADS scores for the biopsied lesions were as follows: 2/3 (66.7%) = PI-RADS 4, 1/3 (33.3%) = PI-RADS 5.

Fig. 2 shows examples of scan1 and scan2 ADC maps calculated from b_0 b_{1500} with a delineated index lesion and normal TZ and PZ regions, and the corresponding ADC histograms illustrating variability between

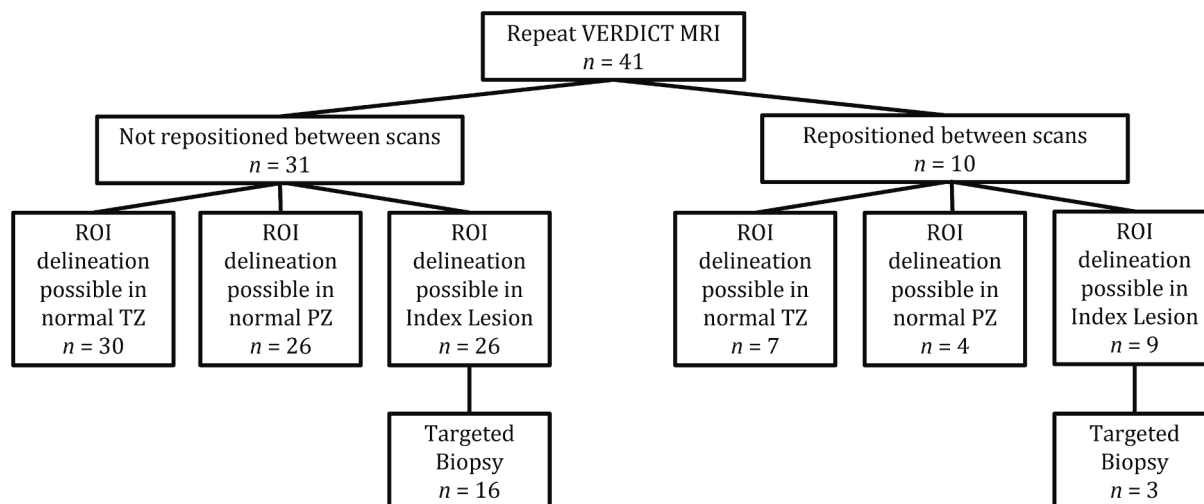


Fig. 1. Flow diagram detailing included participants. VERDICT MRI = Advanced Diffusion MRI sequence (Vascular, Extracellular, Restricted Diffusion for Cytometry in Tumours). Normal TZ = Normal Transition Zone. Normal PZ = Normal Peripheral Zone.

test and retest measurements. Mean ADC and wSD values across all subjects for each of the 5b-value combinations are summarised in Table 4.

For index lesions, ADC showed a moderate negative statistically significant correlation with Likert score when generated from $b_0 b_{1500}$ ($r = -0.52, p < 0.002$) and $b_0 b_{500} b_{2000}$ ($r = -0.56, p < 0.001$), while ADC from $b_0 b_{2000}$ was less correlated with the Likert score ($r = -0.34, p = 0.049$). No statistically significant correlations were observed between Likert score and ADC from $b_0 b_{500}$ ($r = 0.054, p = 0.756$) or $b_{90} b_{500}$ ($r = -0.05, p = 0.77$). The mean ADC values for index lesions were lower than those for normal PZ and similar to normal TZ values.

In a model assessing the effect of ADC magnitude on wSD, no statistically significant relationship was found ($p > 0.05$), thus, estimates of wSD were constant over the range of ADC values, and so wSD was deemed a more appropriate repeatability metric than wCV.

Fig. 3 displays the wSD values and associated confidence intervals for the 5 different b-value combinations of the non-repositioned cohort. For index lesions and normal TZ, there were no statistically significant interactions between b-value combination and whether or not patients had undergone repositioning (index lesions: $p = 0.222$, TZ: $p = 0.061$), therefore the data was pooled. For normal PZ, this interaction was statistically significant ($p = 0.010$) therefore, the 26 patients who were not repositioned and the 4 who were repositioned, were analysed separately.

Across all regions, ADC with $b_{max} = 500$ s/mm² had larger wSD values compared to all other ADC maps. In index lesions, normal TZ, and non-repositioned PZ, wSD was statistically significantly lower for $b_0 b_{1500}$, $b_0 b_{2000}$, and $b_0 b_{500} b_{2000}$, than for $b_{90} b_{500}$ and $b_0 b_{500}$ ($p = 0.001$). For repositioned PZ, wSD was also statistically significantly lower for the $b_0 b_{1500}$ ($p = 0.001$) and $b_0 b_{500} b_{2000}$ ($p = 0.016$) compared to $b_0 b_{500}$. This was also the case for $b_0 b_{1500}$ ($p = 0.011$), $b_0 b_{2000}$ ($p = 0.001$) and $b_0 b_{500} b_{2000}$ ($p = 0.032$) compared to $b_{90} b_{500}$. In all regions wSD was not statistically significantly different between $b_{90} b_{500}$ and $b_0 b_{500}$ ($p > 0.05$).

The wSD of $b_0 b_{1500}$ was statistically significantly lower than $b_0 b_{500} b_{2000}$ in index lesions ($p = 0.001$) and the non-repositioned PZ ($p = 0.001$), however this difference did not reach statistical significance in the TZ ($p = 0.118$) or repositioned PZ ($p = 0.813$). As shown in Table 4, $b_0 b_{1500}$ provided the lowest wSD in all regions in the non-repositioned cohort and for index lesions within the repositioned cohort, while $b_0 b_{2000}$ provided the lowest wSD for normal PZ and TZ in the repositioned cohort. However, the difference between these wSDs never reached statistical significance ($p > 0.05$). Given that $b_0 b_{1500}$ provided the lowest wSD in the majority of regions, this b-value combination was

used for further analysis.

Bland-Altman analysis of ADC from $b_0 b_{1500}$ (Fig. 4) revealed narrow limits of agreement and a consistently small mean difference (bias) of $<0.5 \times 10^{-4}$ mm²/s between ADC values when comparing first and second scans in all regions. Mostly negative bias observed for PZ and index lesions suggested the tendency for lower ADC on the retest scan. (The Bland-Altman plots for other b-value combinations are provided for illustration in Supplementary Figures S1-S4.) Apparent test-retest variability (shown via the Bland-Altman confidence intervals) increased from normal TZ (0.98×10^{-4} mm²/s) versus normal PZ (1.31×10^{-4} mm²/s) versus lesions (1.49×10^{-4} mm²/s). The index lesions from the repositioned group showed a very similar pattern to the index lesions from the non-repositioned group with a mean difference (bias) of -0.42×10^{-4} mm²/s and an apparent test-retest variability of 1.46×10^{-4} mm²/s (Fig. 4d). Additionally, lesion wSD for $b_0 b_{1500}$ did not show a statistically significant correlation to Likert score ($p = 0.95$).

Pooling the data for repositioned ($n = 9$) and non-repositioned ($n = 26$) index lesions (Fig. 4c,d) provided an average wSD of 0.39×10^{-4} mm²/s for $b_0 b_{1500}$. This was closely comparable to the wSD (0.39×10^{-4} mm²/s) of the biopsied cohort with confirmed significant cancer (Gleason Grade $\geq 3 + 4$), which included 3 repositioned and 10 non-repositioned cases (color-coded in Fig. 4c,d). For 7 of the biopsy-confirmed significant cancers, average test-retest ADC was below minimum observed for normal PZ ($<0.32 \times 10^{-3}$ mm²/s).

Pooling the data from all regions, a GLM showed that the size of the ROI was related to the wSD ($p = 0.02$). It can be seen from Fig. 5 that as the number of voxels increases, the log wSD decreases with the adjusted regression coefficient of -1.6×10^{-3} with 95% CIs of $[-3, -0.3] \times 10^{-3}$. This was close to the -1.4×10^{-3} regression value that would be expected for random noise dependence on number of samples ($\frac{1}{\sqrt{\text{VoxelCount}}}$). The ROI color-coding, based on zone shows that index lesions tended to have the lowest voxel counts (clustering below 100 voxels), followed by normal PZ and normal TZ, with the largest region sizes.

4. Discussion

ADC forms part of the mpMRI protocol and has proved valuable in prostate cancer detection and grading [21,22]. The VERDICT protocol has demonstrated utility in prostate cancer grading [16] and also allows generation of ADC surrogate metrics as the sequence consists of multiple b-values with variable diffusion times. Given the dependence on DWI scan parameters, in the current PI-RADS recommendations, which have been rather broad, derived ADC has only been used qualitatively.

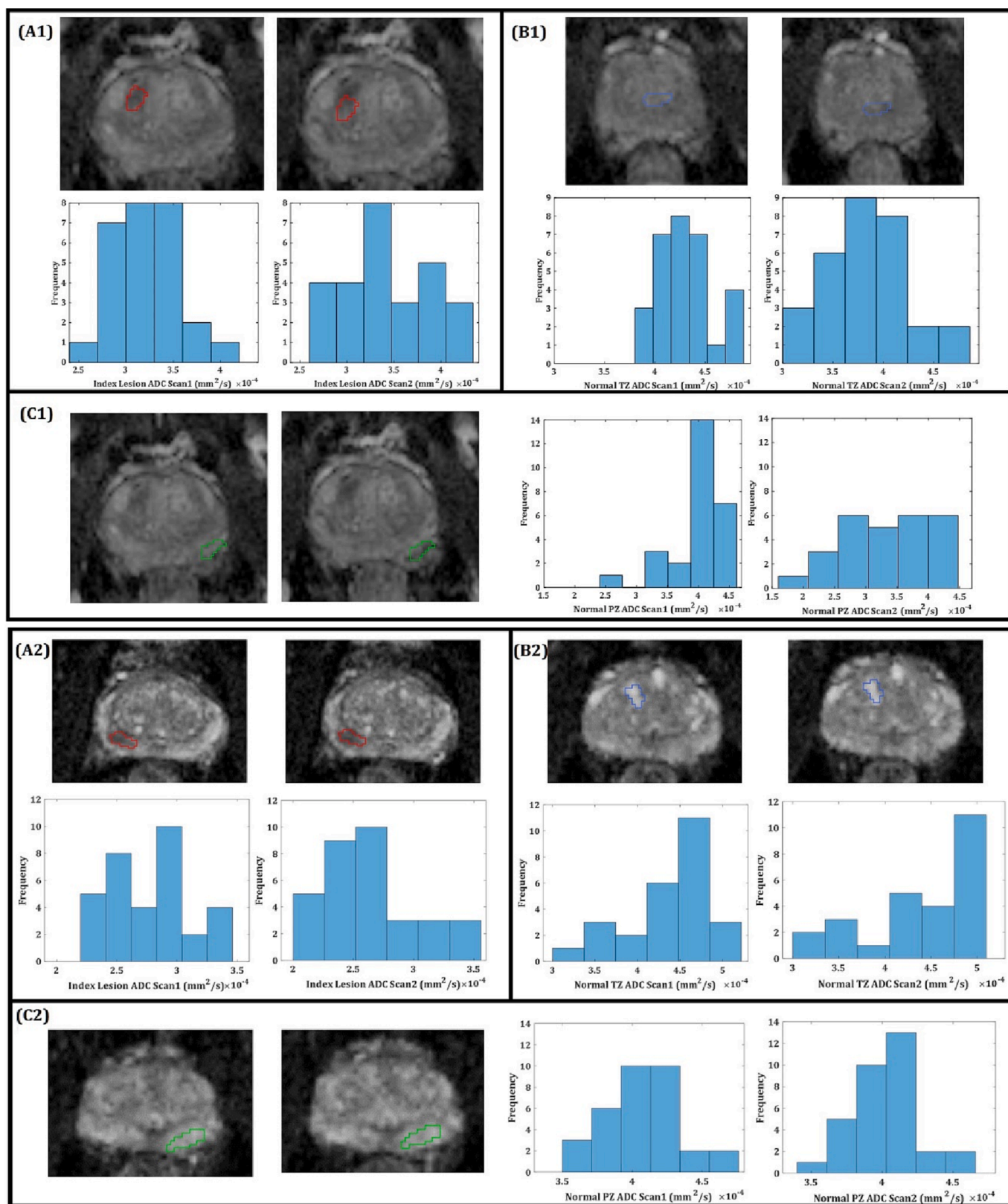


Fig. 2. ADC (Apparent Diffusion Coefficient) maps from two representative patients who was scanned twice back-to-back with no repositioning. ADC maps generated from b_0 b_{1500} s/mm^2 . The panels show the ROIs and corresponding histograms with the distribution of the ADC values for the first and second scans. Top panel index lesion Likert: 4, Gleason Score: 3 + 4. Bottom panel index lesion Likert: 5, Gleason Score: 4 + 3. (A1, A2) Outlined in red is the index lesion (B1, B2) Outlined in blue is an ROI for normal transition zone. (C1, C2) Outlined in green is an ROI for normal peripheral zone. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4

Mean wSD and Mean ADC and Standard Deviations for each *b*-value combination in the different locations for the non-repositioned cohort (n = 31) and the repositioned cohort (n = 10).

<i>b</i> -value Combination s/mm ²	Mean wSD [95% CIs] (x10 ⁻⁴ mm ² /s)	Mean ADC (SD) (x10 ⁻⁴ mm ² /s)	Mean wSD [95% CIs] (x10 ⁻⁴ mm ² /s)	Mean ADC (SD) (x10 ⁻⁴ mm ² /s)	Mean wSD [95% CIs] (x10 ⁻⁴ mm ² /s)	Mean ADC (SD) (x10 ⁻⁴ mm ² /s)
Non-Repositioned						
Normal TZ (n = 30)			Normal PZ (n = 26)		Index Lesion (n = 26)	
0 500	3.08 [2.46, 4.12]	17.14 (3.65)	7.94 [6.25, 10.88]	25.14 (5.77)	8.72 [6.87, 11.95]	18.52 (7.00)
0 1500	0.18 [0.14, 0.24]	3.8 (0.50)	0.53 [0.42, 0.73]	4.46 (0.69)	0.56 [0.44, 0.76]	3.66 (0.68)
0 2000	0.23 [0.18, 0.30]	6.74 (0.80)	0.57 [0.45, 0.79]	7.63 (1.31)	0.83 [0.66, 1.14]	6.12 (1.28)
90 500	1.9 [1.52, 2.54]	20.75 (1.88)	7.15 [5.63, 9.79]	25.22 (3.97)	6.54 [5.15, 8.97]	22.6 (4.59)
0 500 2000	0.34 [0.27, 0.45]	8.32 (1.22)	1.02 [0.80, 1.40]	9.85 (1.72)	1.17 [0.93, 1.61]	7.71 (1.75)
Repositioned						
Normal TZ (n = 7)			Normal PZ (n = 4)		Index Lesion (n = 9)	
0 500	4.5 [2.98, 9.16]	21.14 (4.18)	7.63 [4.57, 21.94]	26.84 (7.22)	10.24 [7.04, 18.69]	23.72 (9.33)
0 1500	0.43 [0.28, 0.87]	4.25 (0.31)	0.4 [0.24, 1.15]	4.54 (0.85)	0.58 [0.40, 1.06]	3.48 (0.68)
0 2000	0.29 [0.19, 0.59]	7.3 (0.73)	0.28 [0.17, 0.82]	7.76 (1.54)	1.79 [1.23, 3.27]	5.78 (1.03)
90 500	3.27 [2.16, 6.65]	22.95 (2.46)	6.44 [3.86, 18.51]	30.13 (5.58)	9.53 [6.56, 17.40]	25.85 (7.29)
0 500 2000	0.83 [0.55, 1.69]	9.32 (0.67)	0.42 [0.25, 1.20]	9.78 (1.84)	1.08 [0.75, 1.98]	7.42 (1.40)

Note. – wSD = within-subject standard deviation, ADC = Apparent Diffusion Coefficient, SD = Standard Deviation, 95% CIs = Confidence Intervals, TZ = Transition Zone, PZ = Peripheral Zone.

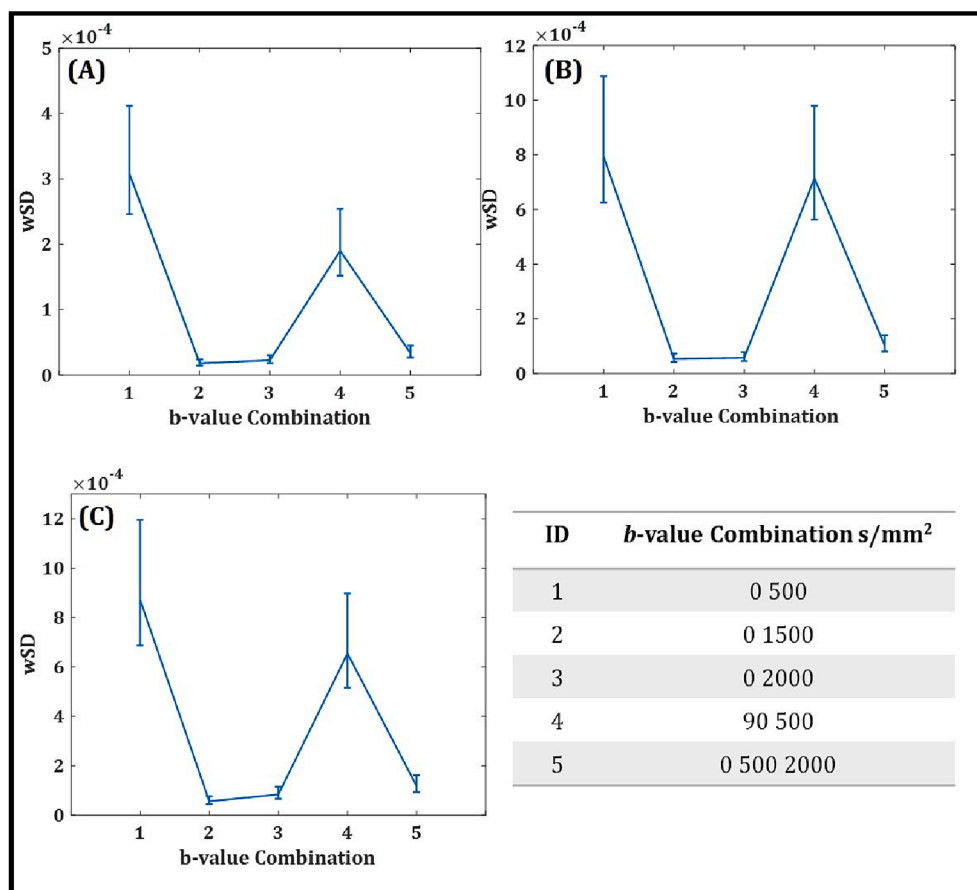


Fig. 3. Wsd (within standard deviation) for each *b*-value combination ADC (Apparent Diffusion Coefficient) map with error bars indicating the confidence intervals for the non-repositioned cohort (A) Normal Transition Zone (B) Normal Peripheral Zone (C) Index Lesion.

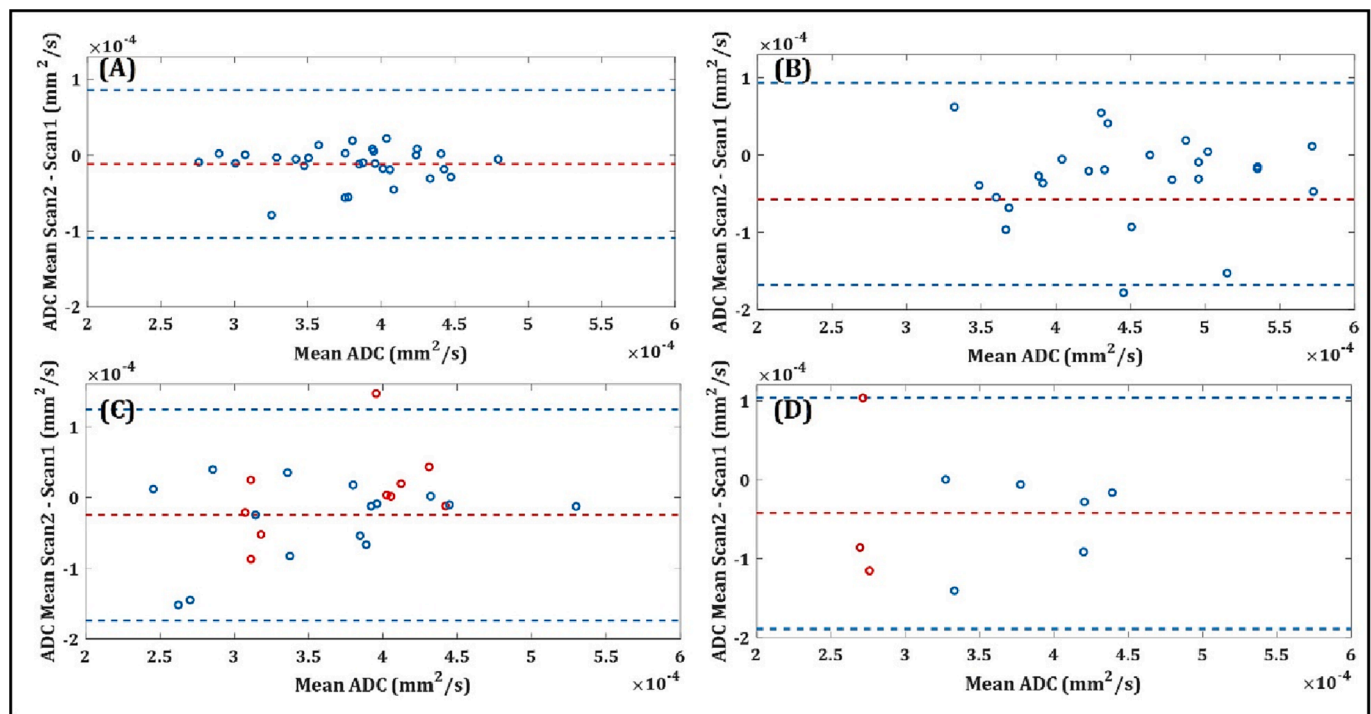


Fig. 4. Bland-Altman analyses showing the level of agreement of mean ADC (Apparent Diffusion Coefficient) between scan 1 and scan 2. ADC generated from $b_0 b_{1500}$ s/mm². Blue dotted lines indicate the limits of agreement (1.96*standard deviation). The Confidence Intervals (CI) for each of the analysed regions are provided: (A) Normal Transition Zone for non-repositioned cohort, (CI: [0.14, 0.24] $\times 10^{-4}$ mm²/s) (B) Normal Peripheral Zone for non-repositioned cohort, (CI: [0.42, 0.73] $\times 10^{-4}$ mm²/s) (C) Index Lesion for non-repositioned cohort (CI: [0.44, 0.76] $\times 10^{-4}$ mm²/s) in red are the index lesions which were significant on biopsy (CI: [0.29, 0.74] $\times 10^{-4}$ mm²/s) (D) Index Lesions for repositioned cohort (CI: [0.40, 1.06] $\times 10^{-4}$ mm²/s), in red are the index lesions which were significant on biopsy (CI: [0.41, 2.70] $\times 10^{-4}$ mm²/s). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

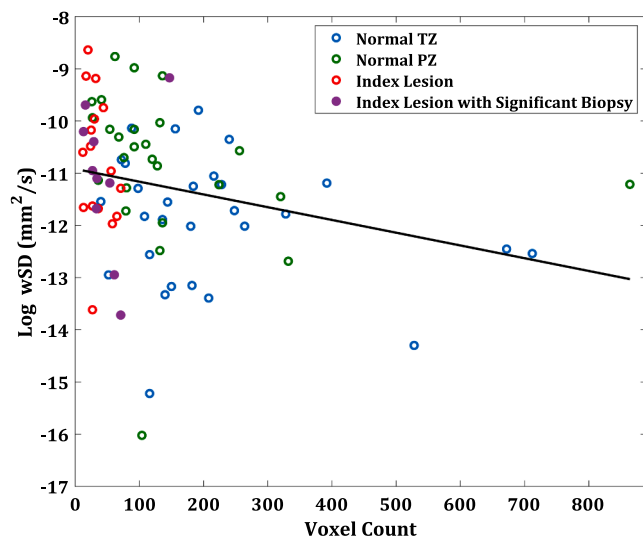


Fig. 5. Pooled data from ROIs from 30 normal transition zone, 26 normal peripheral zone and 26 index lesions (of which 10 had significant cancer confirmed from biopsy) from 31 unique subjects. This demonstrates the relationship between number of voxels within the ROIs and the log wSD (within Standard Deviation) for ADC (Apparent Diffusion Coefficient) maps generated from $b_0 b_{1500}$ s/mm². Generalised estimating equations (solid line) and a Wald test indicated that there was a significant relationship ($p = 0.019$) with the adjusted regression coefficient of -1.6×10^{-3} (CI: $[-3, -0.3] \times 10^{-3}$). The region-specific measurements are colour-coded in the legend.

Recommendations do not include specific guidance on mixing diffusion and echo times outside of utilising $TE \leq 90$ ms. Determining which b -value range provides repeatable ADC is important if the quantitative

value is to be used for discrimination of clinically-significant cancer from benign changes and low-grade disease [3].

This study compared ADC repeatability (quantified by wSD) from different b -value combinations in prostate tissue using the multi b -value VERDICT sequence with variable diffusion times, for subjects who were scanned twice. This study provides the largest repeatability cohort ($n = 41$) published to date, sufficient [13] to allow analysis of repeatability dependence on acquisition parameters, region size and location. Image quality was deemed adequate for all ROIs included in the final analysis. Although VERDICT b -values do not include 800–1000 s/mm² as recommended by PI-RADS, the ADC repeatability trends derived for lower and higher b -values may inform both future repeatability investigations and, ultimately, technical acquisition and analysis recommendations.

Among the five studied b -value combinations, the repeatability was best for ADC from $b_0 b_{1500}$, and was unaffected by ADC magnitude, location, and lesion Likert score. There was a relationship with ROI size; an increase in number of voxels showed a reduction in variability for normal TZ, compared to normal PZ and index lesions (with lowest sizes and highest wSD consistent with increased random noise effect on ADC). Compared to normal PZ and TZ, the observed repeatability was lower (variability higher) for index lesions and marginally better than reported in literature for protocols that used lower b -values and subject repositioning [10–12]. The lack of correlation of ADC repeatability with lesion Likert score for $b_0 b_{1500}$ also suggests that repeatability is independent of disease severity, which would allow establishing technical performance thresholds by pooling data across lesions.

It was demonstrated that ADC with $b_{max} = 500$ s/mm² had the highest variability, indicating low repeatability. Lower repeatability observed for ADC generated from $b_{max} = 500$ s/mm² may be attributed to perfusion effects. At lower b -values the diffusion signal is influenced by fast pseudo-diffusion and pulsatile perfusion effects [23]. The PI-RADS recommendations acknowledge that perfusion information may

be obtained using b -values: 0–500 [24]. The pseudo-diffusion coefficient attributed to this relationship has previously been shown to have poor repeatability and high variability in various organs, including prostate [25–28]. Furthermore, b_{500} may not provide sufficient contrast-to-noise for slow diffusion in prostate tissue at moderately long diffusion times used by the VERDICT protocol.

Larger ROI sizes increased repeatability, possibly due to reduced noise effects on the mean ADC and reduced registration errors between the two scans, as a greater areas and voxel ADC samples were analysed [29]. Therefore, care should be taken when analysing smaller lesions, especially in the PZ where most cancerous lesions are sited.

To reduce kurtosis effects, seen at higher b -values, PI-RADS recommends using $b_{max} = 1000$ for ADC calculations [2]. VERDICT does not include b_{1000} , however, the present study is still in line with QIBA guidelines, which recommend $b_{max} = 500$ –1500 [5]. To reduce variability of multi b -value acquisitions and comply with mono-exponential ADC model assumptions, it is essential to use consistent TEs, which was not possible for retrospective analysis of VERDICT acquisitions in this study. To alleviate inconsistent TE, the studied b -value combinations were limited to those within 15 ms difference. Adhering to PI-RADS recommendations ($TE \leq 90$ ms), shorter diffusion times for high b -values would also help reduce SNR bias in derived mean ADC values.

An intrinsic limitation of this study is that VERDICT uses different TRs and TEs for different b -values which may not be optimised for prostate ADC measurement. Mean ADC values were highly variable across the different b -value combinations. Only b_{2000} has $TR \geq 3,000$ ms, which is the current PI-RADS and QIBA recommendation [2,5]. A longer TR is needed to alleviate T_1 effects for ADC; however, VERDICT is optimised for prostate cancer detection and grading based on the model that quantifies extra- and intra-cellular diffusion fractions, rather than ADC [15]. As VERDICT is clinically promising [16] it was beneficial to investigate whether this sequence could also generate repeatable ADC maps without modifying acquisition parameters, and retrospectively selecting b -value combinations with close TEs.

The lower variability observed for the higher b -value combinations may be related to increased b -value averaging, which is needed due to the decrease in signal with increase in b -value ($b_{90} = 4$ averages, b_0 , b_{500} , b_{1500} and $b_{2000} = 6$ averages). This may not be practical for clinical protocols due to the increased scan time. Additionally, high repeatability does not guarantee clinical diagnostic value. However, utilising $> b_{1000}$ for ADC has shown promise in discrimination between high and low grade prostate tumours [9]. Our results indicated marginal improved correlation of ADC with Likert score for b_0 b_{500} b_{2000} than b_0 b_{1500} , which had better repeatability. The clinical choices of ADC acquisition will need to balance improved stability and improved discriminatory power.

Another limitation, due to the retrospective nature of the analysis, is that the test–retest scans were performed on a single system during a single exam with majority of patients not undergoing repositioning. This only provided short-term ADC repeatability values. These errors could be optimistic for long-term or multi-site prostate ADC studies. However, other research (using lower maximum $b_{max} = 800$ s/mm²) has shown that test–retest repeatability with repositioning and multi-day reproducibility are largely equivalent [30]. Furthermore, acquired data was sufficient for the purpose of the present study to evaluate relative repeatability trends across b -value protocols, prostate zones, region sizes, and ADC values. Additionally, the subset of patients who were repositioned showed similar repeatability results and thus added support for assessed repeatability values. Repeatability of index lesions was similar to that observed for significant cancers confirmed by biopsy.

In summary, ADC maps formed from b_0 b_{1500} combination from the VERDICT acquisition had the least variability. This was found in normal TZ, normal PZ, and prostate cancer lesions. Repeatability increased with ROI size and for normal TZ, versus normal PZ, versus index lesion consistent with reduced noise for mean ADC. The index lesion mean ADC was moderately negatively correlated with Likert score.

Importantly, lesion wSD did not increase with ADC magnitude, hence wSD was a more appropriate repeatability metric than wCV (normalised by mean ADC) for prostate ADC. Future studies should address the limitation of consistent acquisition parameters to optimize ADC acquisitions and evaluate effect of scan resolution for small lesions. Having low variability makes it easier to ascertain whether there have been clinically significant ADC changes in longitudinal therapy response studies and active surveillance prostate imaging programs.

5. Funding Information:

This work was directly supported by Prostate Cancer UK: Targeted Call 2014: Translational Research St.2, project reference PG14-018-TR2. This work was funded in whole or in part with Federal funds from the National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Department of Health and Human Services, under Contract Nos. HHSN268201300071C and HHSN268201500021C.

AB was supported by HEE/NIHR fellowship award HCS SCL-2014-05-002. TC and DM were supported by the NIH/NCI R01CA190299. DM was supported by 75N91021C00036 sub-K. AS-D, TC and DM were supported by U01 CA211205. AS-D was supported by NIH/NCI Cancer Center Support Grant P30 CA008748. SP receives research support from the UK's NIHR UCLH/UCL Biomedical Research Centre.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejrad.2023.110782>.

References

- [1] H.U. Ahmed, A.-E.-S. Bosaily, L.C. Brown, R. Gabe, R. Kaplan, M.K. Parmar, Y. Collaco-Moraes, K. Ward, R.G. Hindley, A. Freeman, A.P. Kirkham, R. Oldroyd, C. Parker, M. Emberton, Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study, *Lancet* 389 (2017) 815–822, [https://doi.org/10.1016/S0140-6736\(16\)32401-1](https://doi.org/10.1016/S0140-6736(16)32401-1).
- [2] B. Turkbey, A.B. Rosenkrantz, M.A. Haider, A.R. Padhani, G. Villeirs, K.J. Macura, C.M. Tempny, P.L. Choyke, F. Cornud, D.J. Margolis, H.C. Thoeny, S. Verma, J. Barentsz, J.C. Weinreb, P.I. Reporting, D.S. Version 2.1, Update of Prostate Imaging Reporting and Data System Version 2, *Eur. Urol.* 76 (2019) (2019) 340–351, <https://doi.org/10.1016/j.eururo.2019.02.033>.
- [3] EAU Guidelines on Prostate Cancer - Uroweb, Uroweb - European Association of Urology. (n.d.). <https://uroweb.org/guidelines/prostate-cancer> (accessed March 21, 2022).
- [4] E.J. Barkovich, P.R. Shankar, A.C. Westphalen, A Systematic Review of the Existing Prostate Imaging Reporting and Data System Version 2 (PI-RADSv2) Literature and Subset Meta-Analysis of PI-RADSv2 Categories Stratified by Gleason Scores, *Am. J. Roentgenol.* 212 (2019) 847–854, <https://doi.org/10.2214/AJR.18.20571>.
- [5] A. Shukla-Dave, N.A. Obuchowski, T.L. Chenevert, S. Jambawalikar, L. H. Schwartz, D. Malyarenko, W. Huang, S.M. Noworolski, R.J. Young, M. S. Shiroishi, H. Kim, C. Coolens, H. Laue, C. Chung, M. Rosen, M. Boss, E. F. Jackson, Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials, *J Magn Reson Imaging.* 49 (2019) e101–e121, <https://doi.org/10.1002/jmri.26518>.
- [6] D.L. Raunig, L.M. McShane, G. Pennello, C. Gatsonis, P.L. Carson, J.T. Voyvodic, R. L. Wahl, B.F. Kurland, A.J. Schwarz, M. Gönen, G. Zahlmann, M.V. Kondratovich, K. O'Donnell, N. Petrick, P.E. Cole, B. Garra, D.C. Sullivan, QIBA Technical Performance Working Group, Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment, *Stat Methods Med Res.* 24 (2015) 27–67, <https://doi.org/10.1177/0962280214537344>.
- [7] C.P. Corona-Villalobos, L. Pan, V.G. Halappa, S. Bonekamp, C.H. Lorenz, J. Eng, I. R. Kamel, Agreement and reproducibility of apparent diffusion coefficient measurements of dual- b -value and multi- b -value diffusion-weighted magnetic resonance imaging at 1.5 Tesla in phantom and in soft tissues of the abdomen, *J Comput Assist Tomogr.* 37 (2013) 46–51, <https://doi.org/10.1097/RCT.0b013e3182720e07>.

- [8] Y. Peng, Y. Jiang, T. Antic, I. Sethi, C. Schmid-Tannwald, S. Eggener, A. Oto, Apparent Diffusion Coefficient for Prostate Cancer Imaging: Impact of b Values, *Am. J. Roentgenol.* 202 (2014) W247–W253, <https://doi.org/10.2214/AJR.13.10917>.
- [9] S.L. Hurrell, S.D. McGarry, A. Kaczmarowski, K.A. Iczkowski, K. Jacobsohn, M. D. Hohenwarter, W.A. Hall, W.A. See, A. Banerjee, D.K. Charles, M.T. Nevalainen, A.C. Mackinnon, P.S. LaViolette, Optimized b-value selection for the discrimination of prostate cancer grades, including the cribriform pattern, using diffusion weighted imaging, *J Med Imaging (Bellingham)*. 5 (2018), 011004, <https://doi.org/10.1117/1.JMI.5.1.011004>.
- [10] P. Gibbs, M.D. Pickles, L.W. Turnbull, Repeatability of echo-planar-based diffusion measurements of the human prostate at 3 T, *Magn Reson Imaging*. 25 (2007) 1423–1429, <https://doi.org/10.1016/j.mri.2007.03.030>.
- [11] G.J.S. Litjens, T. Hambroek, C. Hulsbergen-van de Kaa, J.O. Barentsz, H. J. Huisman, Interpatient variation in normal peripheral zone apparent diffusion coefficient: effect on the prediction of prostate cancer aggressiveness, *Radiology* 265 (2012) 260–266, <https://doi.org/10.1148/radiol.12112374>.
- [12] A. Fedorov, M.G. Vangel, C.M. Tempny, F.M. Fennessy, Multiparametric Magnetic Resonance Imaging of the Prostate, *Invest Radiol.* 52 (2017) 538–546, <https://doi.org/10.1097/RLI.0000000000000382>.
- [13] N.A. Obuchowski, J. Bullen, Quantitative imaging biomarkers: Effect of sample size and bias on confidence interval coverage, *Stat Methods Med Res.* 27 (2018) 3139–3150, <https://doi.org/10.1177/0962280217693662>.
- [14] M.A. Boss, B.S. Snyder, E. Kim, D. Flamini, S. Englander, K.M. Sundaram, N. Gumpeni, S.L. Palmer, H. Choi, A.T. Froemming, T. Persigehl, M.S. Davenport, D. Malyarenko, T.L. Chenevert, M.A. Rosen, Repeatability and Reproducibility Assessment of the Apparent Diffusion Coefficient in the Prostate: A Trial of the ECOG-ACRIN Research Group (ACRIN 6701), *J Magn Reson Imaging*. 56 (2022) 668–679, <https://doi.org/10.1002/jmri.28093>.
- [15] C. Bailey, R.M. Bourne, B. Siow, E.W. Johnston, M. Brizmohun Appayya, H. Pye, S. Heavey, T. Mertzaniidou, H. Whitaker, A. Freeman, D. Patel, G.L. Shaw, A. Sridhar, D.J. Hawkes, S. Punwani, D.C. Alexander, E. Panagiotaki, VERDICT MRI validation in fresh and fixed prostate specimens using patient-specific moulds for histological and MR alignment, *NMR Biomed.* 32 (2019) e4073.
- [16] E.W. Johnston, E. Bonet-Carne, U. Ferizi, B. Yvernault, H. Pye, D. Patel, J. Clemente, W. Piga, S. Heavey, H.S. Sidhu, F. Giganti, J. O'Callaghan, M. Brizmohun Appayya, A. Grey, A. Saborowska, S. Ourselin, D. Hawkes, C. M. Moore, M. Emberton, H.U. Ahmed, H. Whitaker, M. Rodriguez-Justo, A. Freeman, D. Atkinson, D. Alexander, E. Panagiotaki, S. Punwani, VERDICT MRI for Prostate Cancer: Intracellular Volume Fraction versus Apparent Diffusion Coefficient, *Radiology* 291 (2019) 391–397, <https://doi.org/10.1148/radiol.2019181749>.
- [17] S. Singh, H. Rogers, B. Kanber, J. Clemente, H. Pye, E.W. Johnston, T. Parry, A. Grey, E. Dinneen, G. Shaw, S. Heavey, U. Stopka-Farooqui, A. Haider, A. Freeman, F. Giganti, D. Atkinson, C.M. Moore, H.C. Whitaker, D.C. Alexander, E. Panagiotaki, S. Punwani, Avoiding Unnecessary Biopsy after Multiparametric Prostate MRI with VERDICT Analysis: The INNOVATE Study, *Radiology* 305 (2022) 623–630, <https://doi.org/10.1148/radiol.212536>.
- [18] E. Johnston, H. Pye, E. Bonet-Carne, E. Panagiotaki, D. Patel, M. Galazi, S. Heavey, L. Carmona, A. Freeman, G. Trevisan, C. Allen, A. Kirkham, K. Burling, N. Stevens, D. Hawkes, M. Emberton, C. Moore, H.U. Ahmed, D. Atkinson, M. Rodriguez-Justo, T. Ng, D. Alexander, H. Whitaker, S. Punwani, INNOVATE: A prospective cohort study combining serum and urinary biomarkers with novel diffusion-weighted magnetic resonance imaging for the prediction and characterization of prostate cancer, *BMC Cancer* 16 (2016) 816, <https://doi.org/10.1186/s12885-016-2856-2>.
- [19] J.I. Epstein, L. Egevad, M.B. Amin, B. Delahunt, J.R. Srigley, P.A. Humphrey, Grading Committee, The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System, *Am J Surg Pathol.* 40 (2016) 244–252. Doi: 10.1097/PAS.0000000000000530.
- [20] A. El-Shater Bosaily, C. Parker, L.C. Brown, R. Gabe, R.G. Hindley, R. Kaplan, M. Emberton, H.U. Ahmed, PROMIS — Prostate MR imaging study: A paired validating cohort study evaluating the role of multi-parametric MRI in men with clinical suspicion of prostate cancer, *Contemp Clin Trials*. 42 (2015) 26–40, <https://doi.org/10.1016/j.cct.2015.02.008>.
- [21] J.H. Kim, J.K. Kim, B.-W. Park, N. Kim, K.-S. Cho, Apparent diffusion coefficient: Prostate cancer versus noncancerous tissue according to anatomical region, *J. Magn. Reson. Imaging* 28 (2008) 1173–1179, <https://doi.org/10.1002/jmri.21513>.
- [22] G.J.S. Litjens, R. Elliott, N.N.C. Shih, M.D. Feldman, T. Kobus, C. Hulsbergen-van de Kaa, J.O. Barentsz, H.J. Huisman, A. Madabhushi, Computer-extracted Features Can Distinguish Noncancerous Confounding Disease from Prostatic Adenocarcinoma at Multiparametric MR Imaging, *Radiology* 278 (2016) 135–145, <https://doi.org/10.1148/radiol.2015142856>.
- [23] D. Le Bihan, E. Breton, D. Lallemand, M.L. Aubin, J. Vignaud, M. Laval-Jeantet, Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging, *Radiology* 168 (1988) 497–505, <https://doi.org/10.1148/radiology.168.2.3393671>.
- [24] J.C. Weinreb, J.O. Barentsz, P.L. Choyke, F. Cornud, M.A. Haider, K.J. Macura, D. Margolis, M.D. Schnall, F. Shtern, C.M. Tempny, H.C. Thoeny, S. Verma, PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2, *Eur Urol.* 69 (2016) 16–40. Doi: 10.1016/j.eururo.2015.08.052.
- [25] Y. Cui, H. Dyvorne, C. Besa, N. Cooper, B. Taouli, IVIM diffusion-weighted imaging of the liver at 3.0 T: Comparison with 1.5 T, *Eur J Radiol Open.* 2 (2015) 123–128, <https://doi.org/10.1016/j.ejro.2015.08.001>.
- [26] A. Jakab, R. Tuura, R. Kottke, C.J. Kellenberger, I. Scheer, Intra-voxel incoherent motion MRI of the living human foetus: technique and test-retest repeatability, *Eur Radiol Exp.* 1 (2017), <https://doi.org/10.1186/s41747-017-0031-4>.
- [27] N.P. Jerome, K. Miyazaki, D.J. Collins, M.R. Orton, J.A. d'Arcy, T. Wallace, L. Moreno, A.D.J. Pearson, L.V. Marshall, F. Carceller, M.O. Leach, S. Zacharoulis, D.-M. Koh, Repeatability of derived parameters from histograms following non-Gaussian diffusion modelling of diffusion-weighted imaging in a paediatric oncological cohort, *Eur Radiol.* 27 (2017) 345–353, <https://doi.org/10.1007/s00330-016-4318-2>.
- [28] C. Reischauer, R. Patzwahl, D.-M. Koh, J.M. Froehlich, A. Gutzeit, Non-Mono-Exponential Analysis of Diffusion-Weighted Imaging for Treatment Monitoring in Prostate Cancer Bone Metastases, *Sci Rep.* 7 (2017), <https://doi.org/10.1038/s41598-017-06246-4>.
- [29] M.E. Miquel, A.D. Scott, N.D. Macdougall, R. Boubertakh, N. Bharwani, A. G. Rockall, In vitro and in vivo repeatability of abdominal diffusion-weighted MRI, *Br J Radiol.* 85 (2012) 1507–1512, <https://doi.org/10.1259/bjr/32269440>.
- [30] M.A. Boss, B.S. Snyder, E. Kim, D. Flamini, S. Englander, K.M. Sundaram, N. Gumpeni, S.L. Palmer, H. Choi, A.T. Froemming, T. Persigehl, M.S. Davenport, D. Malyarenko, T.L. Chenevert, M.A. Rosen, Repeatability and Reproducibility Assessment of the Apparent Diffusion Coefficient in the Prostate: A Trial of the ECOG-ACRIN Research Group (ACRIN 6701), *Journal of Magnetic Resonance Imaging*. n/a (n.d.). Doi: 10.1002/jmri.28093.