札幌医学雑誌　**86**（1 － 6）85 ～ 93（2017）　doi：10.15114 / smj.86.85

ORIGINAL

Are ADC histogram metrics repeatable?

Koichi ONODERA[1)], Naoya YAMA[1)], Maki ONODERA[1)], Naomi KOYAMA[1)], Yoshifusa KYUNA[1)], Mitsuhiro NAKANISHI[2)], Masamitsu HATAKENAKA[1)]

[1)] Sapporo Medical University, Diagnostic Radiology

[2)] Sapporo Medical University Hospital, Division of Radiology

ABSTRACT

Purpose: The present study evaluated the repeatability of apparent diffusion coefficient (ADC) histogram metrics in clinical MRI.

Methods: Twelve patients who underwent head MRI in our hospital from May to July in 2016 were included in the present study. All patients gave informed consent. Two sequential diffusion-weighted images with echo planar imaging (DWI-EPI) in the identical positioning were obtained. The b-factors of 0 and 1000 or 1500 s/mm^2 were used, three orthogonal motion proving gradients (MPGs) were applied, and synthesized images were generated. The regions of interest (ROIs) were assigned at the lesions on the 1st DWI and pasted onto the 2nd at the same size and location. Voxel-wise ADC was calculated by fitting the signal intensity change of each voxel into a mono-exponential curve. ADCs calculated from 1st and 2nd DWI were defined as ADC-1st and ADC-2nd, respectively. To investigate the repeatability of voxel-wise ADC in each lesion, ADC-1st and ADC-2nd were compared using Wilcoxon matched-pairs signed rank test and linear regression. To consider repeatability of ADC histogram metrics for all lesions, minimal, 25%, median, 75%, maximum, mean, skewness, and kurtosis of ADC-1st and ADC-2nd for each lesion were compared using linear regression and Bland-Altman plot.

Results: For repeatability of voxel-wise ADC, significant differences were observed between ADC-1st and ADC-2nd in 5 lesions. Linear regression did not show significance of the slope in 5 lesions. As for repeatability of ADC histogram metrics, all ADC histogram metrics except skewness and kurtosis showed significance of the slope in linear regression ($p<0.0001$) and high repeatability in Bland-Altman plot.

Conclusion: The histogram metrics of voxel-wise ADC like minimum, 25%, median, 75%, maximum, and mean show high repeatability, but skewness and kurtosis did not.

(Accepted September 28, 2017)

Key words: DWI, ADC, histogram, repeatability

1 Introduction

The clinical utility of diffusion-weighted imaging (DWI) is widely acknowledged, not only in diagnosing acute brain infarction but also for differentiating malignancy from benignancy,[1)] evaluating malignant potential of the tumor, and predicting prognosis.[2, 3)] Studies reporting that apparent diffusion coefficient (ADC) histogram metrics are useful for evaluating malignant potential of the tumor have recently increased.[4-7)] However, different results have also been reported.[8, 9)] DWI using echo-planar imaging (DWI-EPI) has a distortion artifact, and the repeatability of ADC histogram metrics have not yet been evaluated sufficiently. The purpose of the present study is to evaluate repeatability of voxel-wise ADC histogram metrics in clinical DWI.

2 Materials and Methods

2・1・1 Clinical study

Seventeen consecutive patients who underwent clinical head MRI using whole-body clinical MRI systems and head coil (1.5 and 3T of Ingenia, Philips Medical Systems, Best, The Netherlands and 3T of Signa HDx , GE Healthcare, Milwaukee, WI, USA, respectively) in our hospital from May to July in 2016 were included in the present study. All patients gave informed consent to participate in the study. Five patients were excluded for analysis due

to poor image quality or post-operative state of the tumor. Thus, the present study included 12 patients. Two sequential free-breathing DWI-EPI with identical positioning were obtained during MRI examination. The b-factors of 0 and 1000 or 1500 s/mm^2 were used, and three orthogonal motion proving gradients (MPGs) were applied, and synthesized images were generated. The number of excitation (NEX) was one. The regions of interest (ROIs) were assigned at the lesions on the 1st DWI and pasted onto the 2nd at the same size and location. Voxel-wise ADC was calculated by fitting signal intensity change of each voxel into a mono-exponential curve. ADCs calculated from 1st and 2nd DWI were referred to as ADC-1st and ADC-2nd, respectively. Details of the clinical information and MRI parameters are presented in Table 1.

2・1・2 Phantom study

ADC and T2 of the phantom containing 0.2 mM gadolinium and 80% polyvinyl alcohol (Nikkofines, 90-401 type, Tokyo, Japan) were measured in a whole-body 3T clinical MRI system using a head coil (Ingenia, Philips Medical Systems, Best, The Netherlands). To calculate T2, spin echo images (repetition time=2000 ms, echo time=10, 20, 30, 40, 50, 60, 70, and 80 ms, slice thickness=4 mm, field of view (FOV)=220 mm, matrix=112*112) with one and two NEX were obtained two times sequentially with identical positioning. The ROI containing 100 voxels was assigned at the center of the phantom, and each voxel T2 was calculated by fitting signal intensity at echo times of 10 and 80 ms into a mono-exponential curve. The ROI for the 1st examination image was copied and pasted onto the 2nd at the same size and location.

To calculate ADC, DWI-EPI and DWI using turbo spin-echo (DWI-TSE) (repetition time=4000 ms, echo time=85 ms for EPI and 87.4 ms for TSE, b-factor =0 and 1000 s/mm^2, synthesized image from three orthogonal MPGs, slice thickness=4 mm, field of view (FOV)=220 mm, matrix=112*110 for EPI and 112*112 for TSE) with three and 18 NEX for DWI-EPI and 16 NEX for DWI-TSE were obtained two times sequentially with an identical positioning. These NEX (18 and 16) are the maximum for DWI-EPI and DWI-TSE in the MRI system used. The ROI containing 100 voxels was assigned at the center of the phantom, and voxel-wise ADC was calculated by fitting signal intensity into a mono-exponential curve. Again, the ROI for the 1st examination image was copied and pasted onto the 2nd one holding its size and location.

2・2 Statistics

Statistical analyses were performed by using the GraphPad Prism version 7.00 (GraphPad Software, Inc., San Diego, CA, USA) and SPSS Statistics version 24 (IBM Corp., Armonk, NY, USA). $p<0.05$ was considered statistically significant.

Table 1. Clinical information and MRI parameters

Case No.	Age	Gender	Clinical/histological diagnosis	No. of voxels in the ROI	Vender	Magnetic field strength (T)	slice thickness (mm)	FOV (mm)	Matrix	TR (ms)	TE (ms)	b value (s/mm^2)
1	77	M	meningioma	132	PHILIPS	3	3	200	128*154	4922.4	95.5	1500
2	85	F	metastatic bone tumor	156	GE	3	4	200	128*192	10250	80.7	1500
3	60	F	meningioma	110	GE	3	3	200	128*192	5500	80.9	1500
4	86	F	pituitary adenoma	154	GE	3	3	200	128*192	5500	80.8	1500
5	77	F	colloid plexus tumor	56	PHILIPS	1.5	4	200	112*127	4000	90	1000
6	33	F	oligodendroglioma+	144	GE	3	3	200	128*192	10000	80.9	1500
7	69	M	meningioma	144	PHILIPS	1.5	4	220	112*130	4279.1	90	1000
8	36	M	central neurocytoma+	49	GE	3	3	200	128*192	10450	80.9	1500
9	32	F	meningothelial meningioma+	35	GE	3	3	200	128*192	10000	80.8	1500
10	77	F	meningioma	111	GE	3	3	200	128*192	9525	80.8	1500
11	69	F	diffuse large B cell lymphoma+	79	GE	3	3	200	128*192	10475	81.4	1500
12	40	M	anaplastic oligodendroglioma+	56	GE	3	3	210	128*192	10300	80.6	1500

(+) means histological diagnosis. ROI, FOV, TR, and TE mean region of interest, field of view, repetition time, and echo time, respectively.

2 · 2 · 1 Clinical study

To evaluate repeatability of voxel-wise ADC in each lesion, the distribution of voxel-wise ADC was analyzed by D'Agostino & Pearson normality test. The ADC-1st and ADC-2nd were compared using Wilcoxon matched-pairs signed rank test and linear regression. For repeatability of ADC histogram metrics for all lesions, the minimal, 25%, median, 75%, maximum, mean, skewness, and kurtosis of ADC-1st and ADC-2nd for each lesion were analyzed by D'Agostino & Pearson normality test and compared using linear regression and Bland-Altman plot.

2 · 2 · 2 Phantom study

The repeatability of T2 and ADC was analyzed by Bland-Altman plot.

3 Results

3 · 1 Clinical study

The voxel-wise ADC did not show normal distribution in 5 lesions (D'Agostino & Pearson normality test). Significant differences were observed between ADC-1st and ADC-2nd in 5 lesions (Wilcoxon matched-pairs signed rank test). Linear regression did not show significance of the slope in 5 lesions. Detail results are summarized in Table 2, and representative cases are demonstrated in Figs. 1 and 2.

As for repeatability of ADC histogram metrics, the minimal, 25%, median, 75%, maximum, mean and skewness of ADC showed normal distribution but kurtosis did not (D'Agostino & Pearson normality test). The minimal, 25%, median, 75%, maximum, and mean of ADC showed significance of the slope in linear regression ($p<0.0001$), but skewness or kurtosis showed no significance. All histogram metrics except skewness and kurtosis showed high repeatability in Bland-Altman plot. Results are summarized in Fig. 3 and Table 3.

Table 2. Repeatability of voxel-wise ADC within ROI

Case No.	D'Agostino & Pearson normality test (p)	Wilcoxon matched-pairs signed rank test (p)	Linear regression		
			Significance of slope	r2	Equation
1	<0.05	-	<0.0001	0.266	Y=0.5109X+0.4183
2	-	-	<0.0001	0.729	Y=0.8883X+0.1004
3	<0.05	<0.0001	-	0.004	Y=0.05564X+0.6417
4	-	-	-	0.017	Y=0.1063X+0.3934
5	-	0.02	<0.0001	0.545	Y=0.7857X+0.3176
6	<0.05	-	-	<0.001	Y=-0.00135X+1.352
7	-	0.0003	<0.0001	0.268	Y=0.4401X+0.5292
8	-	-	0.003	0.176	Y=0.5665X+0.2496
9	-	0.02	-	0.084	Y=0.3408X+0.5238
10	<0.05	-	-	<0.001	Y=0.01436X+0.7364
11	<0.05	-	<0.001	0.337	Y=0.7951X+0.1569
12	-	0.003	<0.0001	0.270	Y=0.5868X+0.5514

(-) means p>0.05.

Table 3. Linear regression and Bland-Altman plot parameters for ADC histogram metrics

ADC histogram metrics.	Linear regression			Bland-Altman plot (% difference)		
	Significance of slope	r2	Equation	Bias	Standard deviation of bias	95% confidence interval
minimum	<0.0001	0.9871	Y=1.0469X-0.04821	-3.867	6.643	-16.89 to 9.153
25%	<0.0001	0.9953	Y=1.027X-0.03676	-2.031	3.34	-8.567 to 4.516
median	<0.0001	0.9946	Y=0.999X-0.01243	1.521	3.255	-4.86 to 7.901
75%	<0.0001	0.990	Y=1.012X-0.01846	-0.8745	3.991	-8.697 to 6.948
maximum	<0.0001	0.9358	Y=0.986X+0.04819	2.617	9.155	-15.33 to 20.56
mean	<0.0001	0.9945	Y=-1.01X-0.01911	1.269	3.044	-4.697 to 7.236
skewness	-	0.003	Y=0.0631X+0.2921	3.089	227.5	-442.8 to 449
kurtosis	-	0.006	Y=0.1208X-0.01962	187.5	837	(-1453 to 1828)

(-) means p>0.05.

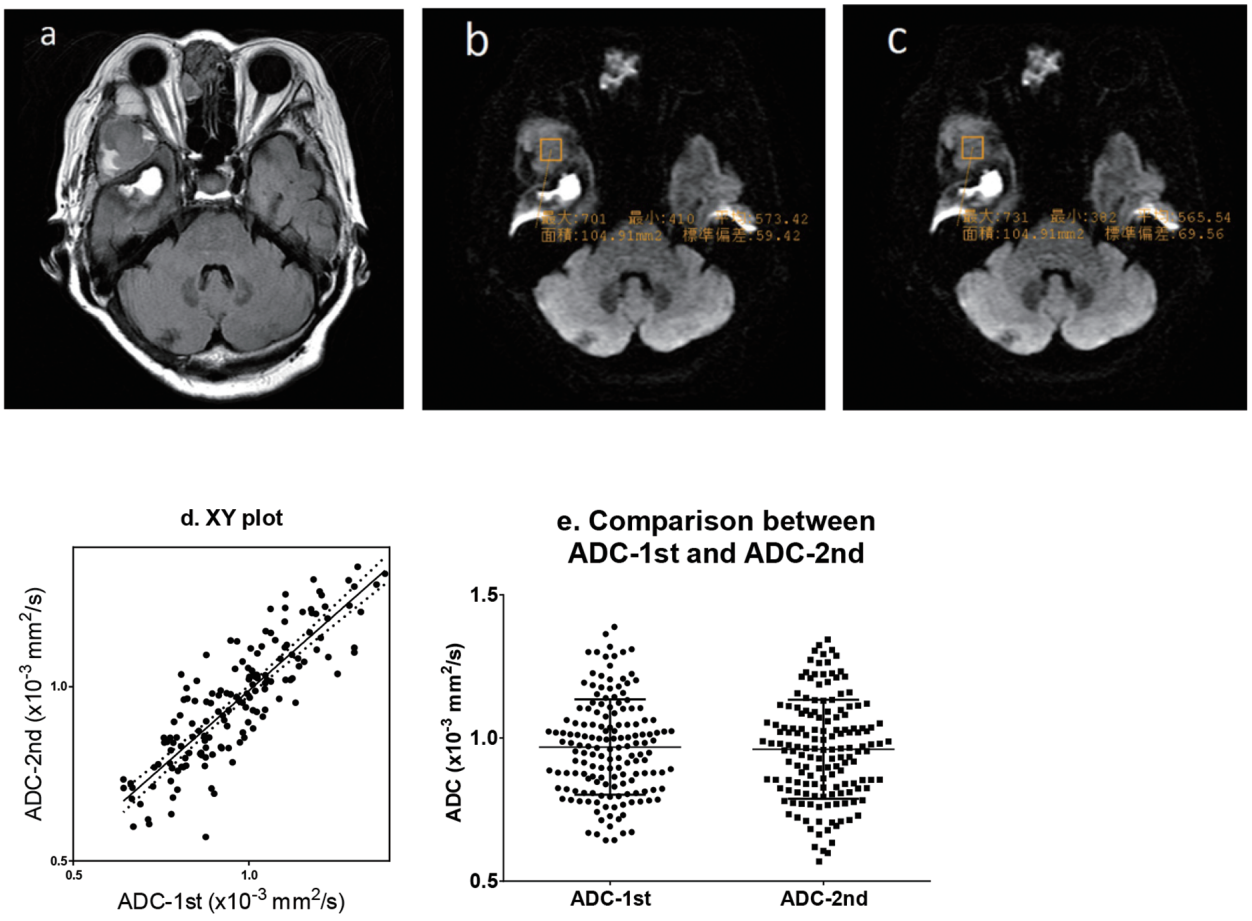

Fig. 1. Good repeatability case (case 2: metastatic bone tumor)
(a), (b), and (c) show fluid-attenuated inversion-recovery image, 1st DWI (b=1500 s/mm^2), and 2nd DWI (b=1500 s/mm^2), respectively. (d) shows XY plot between ADC-1st and ADC-2nd; line and dotted lines indicate regression line and 95% confidence intervals. (e) shows scatter plot of ADC-1st and ADC-2nd; three horizontal lines represent 25%, 50%, and 75% values, respectively.
It is interesting that the lesion showed high repeatability despite being in a region prone to distortion artifact.

3・2 Phantom study

The 95% confidence interval of T2 was smaller than that of ADC from DWI-EPI and DWI-TSE. The 95% confidence interval decreased with increasing NEX for both T2 and ADC. The data range from minimum to maximum also narrowed with increasing NEX for ADC, but was not evident for T2. Details are summarized in Fig. 4 and Table 4.

4 Discussion

The repeatability of voxel-wise ADC was not high in the present study. We expected much higher repeatability due to ideal study conditions. Surprisingly, 5 out of 12 lesions showed significant difference between ADC-1st and ADC-2nd, in addition, 5 out of 12 lesions showed no significance of the slope in linear regression. These results suggest that voxel-wise ADC is not repeatable sufficiently.

In contrast, as for ADC histogram metrics all metrics except skewness and kurtosis showed high repeatability. While this may appear strange, but the finding that minimal, 25%, median, 75%, maximum, and mean of ADC held to some extent even in case 6 that showed poor repeatability supports the results. The voxel of the minimum ADC-1st shows almost median value in the 2nd DWI, but the value of minimum ADC-2nd is almost the same (within 1%) as minimum ADC-1st even though the voxels showing minimum value are different in the 1st and 2nd DWIs (Fig. 2). We consider that the ADC histogram metrics like minimal, 25%, median, 75%, maximum, and mean of ADC are robust to data variability.

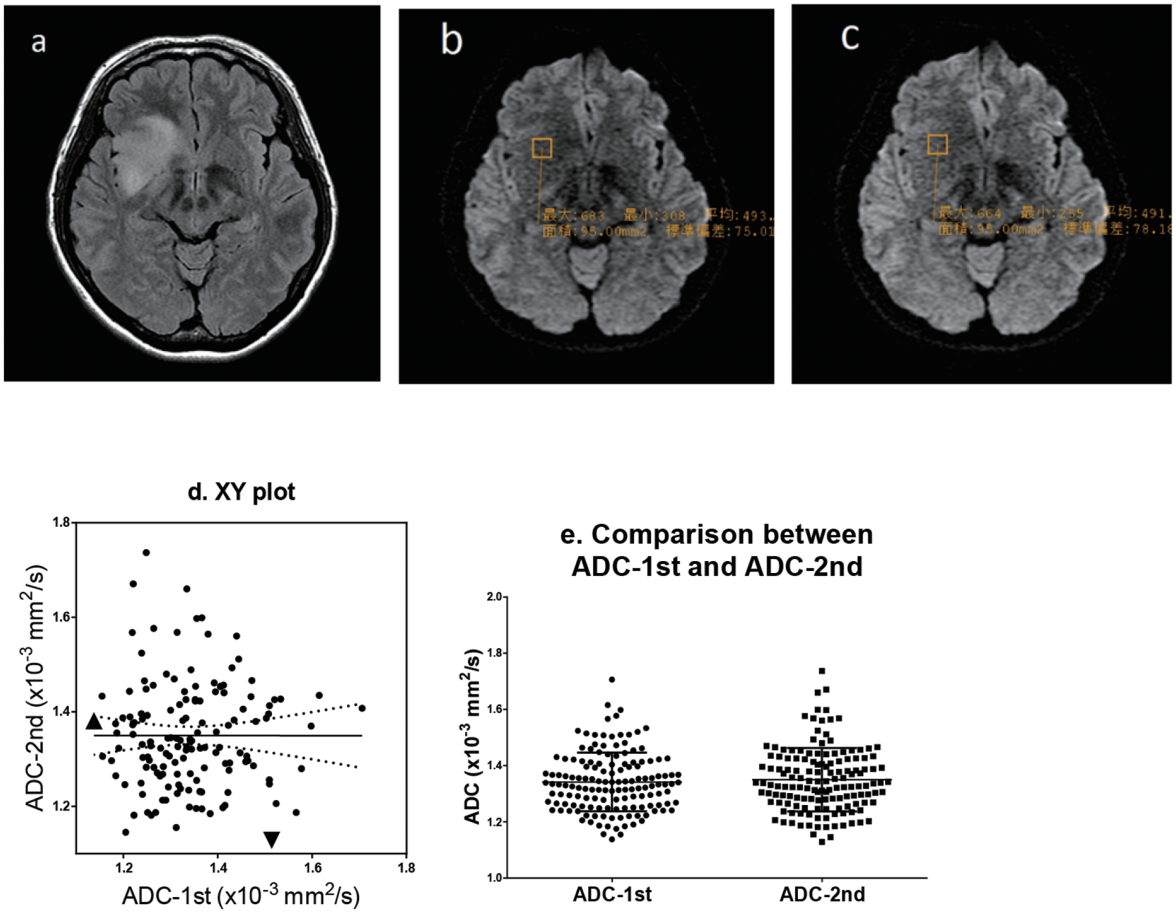

Fig. 2. Poor repeatability case (case 6: oligodendroglioma)
(a), (b), and (c) show fluid-attenuated inversion-recovery image, 1st DWI (b=1500 s/mm^2), and 2nd DWI (b=1500 s/mm^2), respectively. (d) shows XY plot between ADC-1st and ADC-2nd; line and dotted lines indicate regression line and 95% confidence intervals. The voxel showing lowest ADC-1st (▲, 1.138*10^{-3} mm^2/s) is different from that showing lowest ADC-2nd (▼, 1.128*10^{-3} mm^2/s) but the difference (100*(1.138-1.128)/1.138=0.879) was within 1%. (e) shows scatter plot of ADC-1st and ADC-2nd; three horizontal lines indicate 25%, 50%, and 75% values, respectively.

Table 4. Bland-Altman plot parameters for T2 and ADC in phantom study

Measured parameter	Number of excitation	Examination time (s)	Bland-Altman plot (% difference)		
			Bias	Standard deviation of bias	95% confidence interval
T2	1	226	-0.0546	0.5927	-1.216 to 1.107
	2	450	0.1761	0.4481	-0.702 to 1.054
ADC from DWI- EPI	3	56	0.2811	4.967	-9.455 to 10.02
	18	292	0.5665	1.741	-2.845 to 3.978
ADC from DWI- TSE	3	48	3.524	6.875	-9.951 to 17.00
	16	288	-1.220	2.430	-5.984 to 3.544

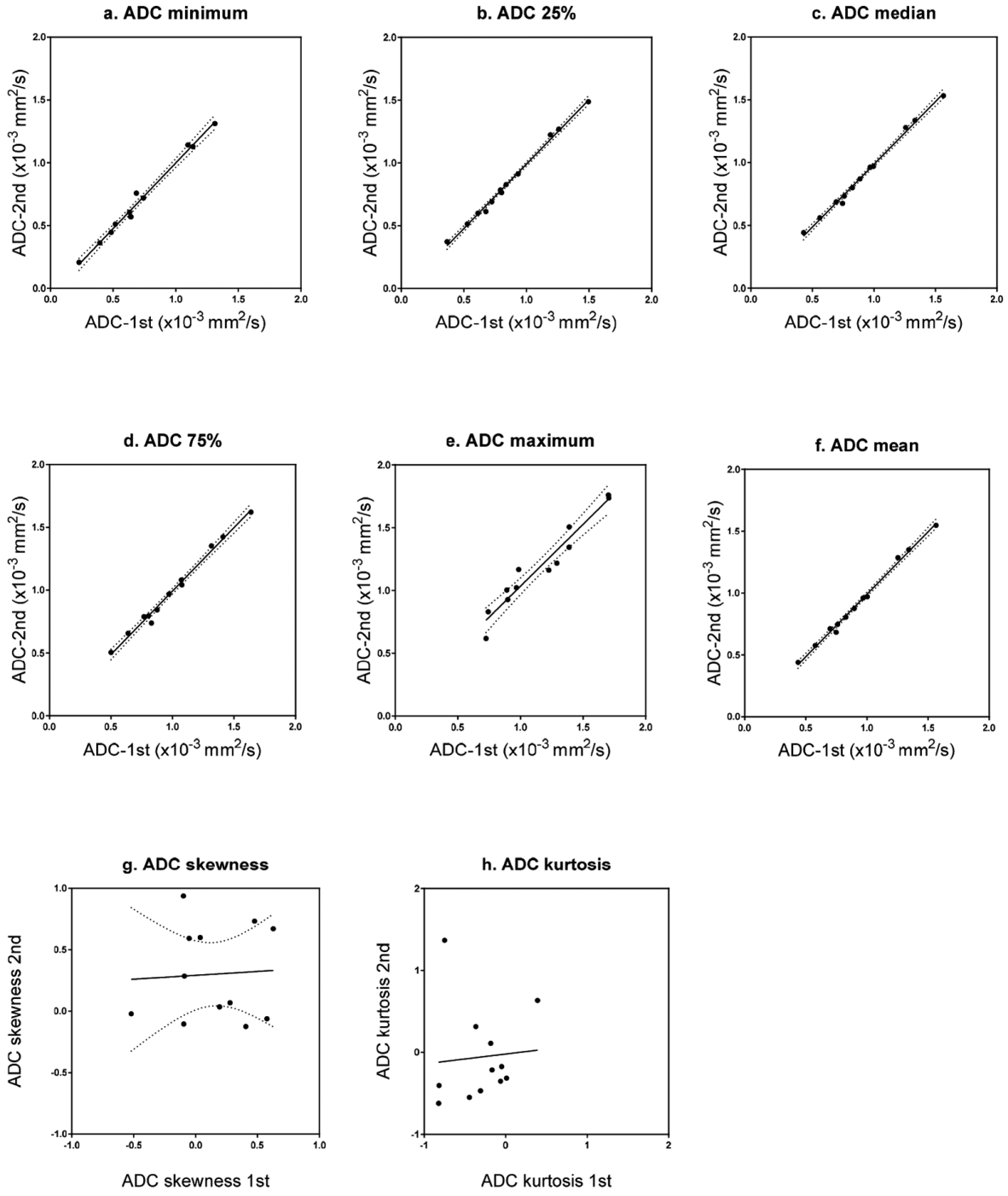

Fig. 3. XY plots of ADC histogram metrics

(a), (b), (c), (d), (e), (f), (g), and (h) show XY plots of minimal, 25%, median, 75%, maximum, mean, skewness, and kurtosis of ADC-1st and ADC-2nd, respectively; line and dotted lines indicate regression line and 95% confidence intervals. ADC kurtosis did not show normal distribution, thus 95% confidence interval line was not drawn.

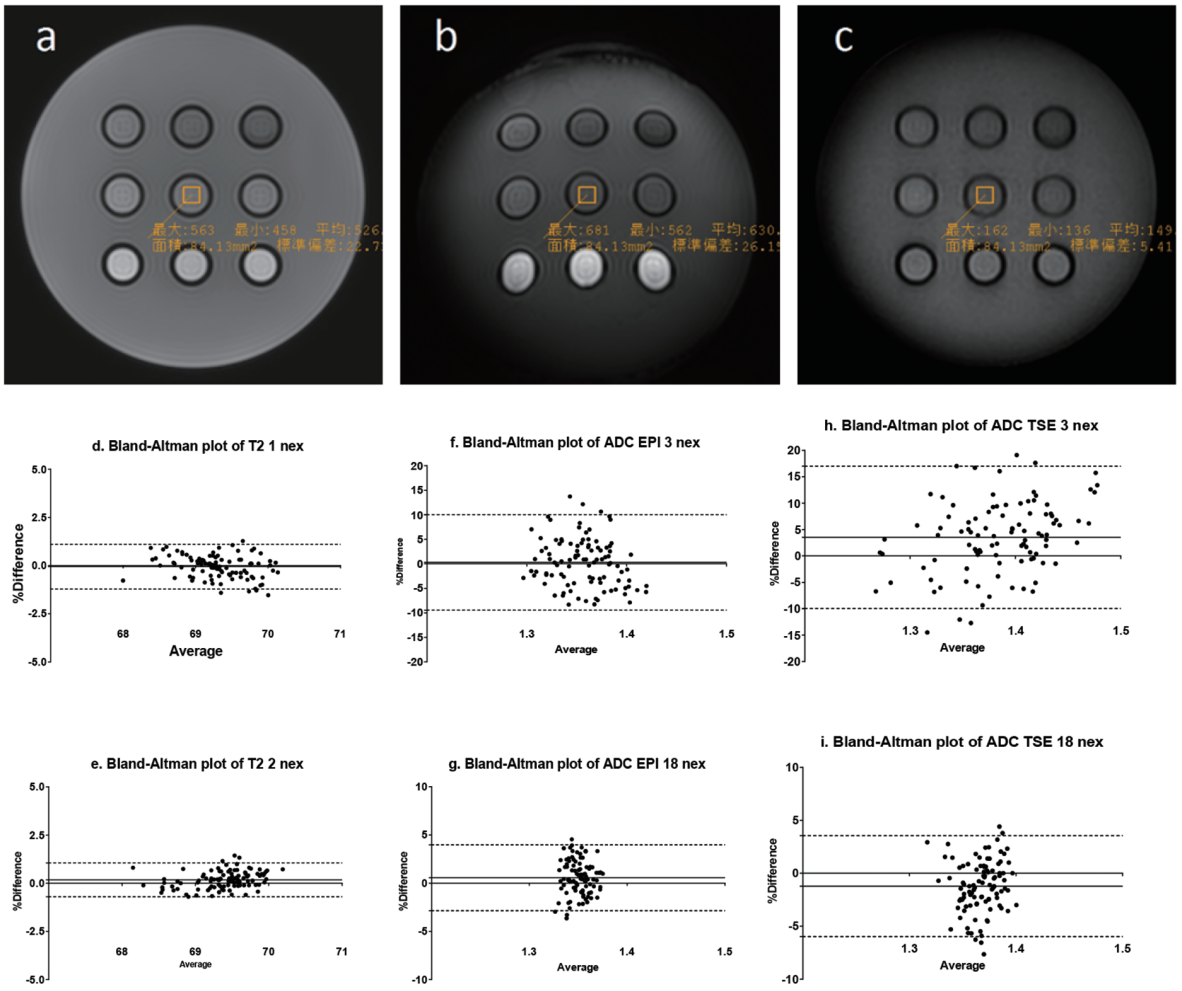

Fig. 4. Phantom images and Bland–Altman plot of phantom data
(a), (b), and (c) show spin echo image (echo time =80 ms, two NEX), DWI-EPI (b=1000 s/mm^2, 18 NEX), and DWI-TSE (b=1000 s/mm^2, 16 NEX), respectively. (d) and (e) show T2 with one and two NEX, respectively. (f) and (g) show ADC obtained from DWI-EPI with three and 18 NEX, respectively. (h) and (i) show ADC obtained from DWI-TSE with three and 16 NEX, respectively.

ADC skewness and kurtosis showed low repeatability. This result is probably due to the low repeatability of voxel-wise ADC itself. Radiologists should be aware of this characteristic and therefore pay sufficient attention in the interpretation of DWI and related metrics like skewness and kurtosis. The conflicting results have been reported regarding correlation between ADC skewness or kurtosis, and malignant potential of the prostate lesion.[8, 9] The low repeatability of ADC skewness and kurtosis would be the reason for the discrepancy.

To investigate the reasons for low repeatability of voxel-wise ADC, phantom studies were performed. Voxel-wise ADC showed low repeatability compared with voxel-wise T2 irrespective of DWI-EPI or DWI-TSE. We interpret the results as follows: the sequence of DWI is more complex than that of spin echo imaging because MPG and readout gradient for EPI or TSE were added and these complexed sequences could lead to increase in data variance. Distortion artifacts related to EPI might be one of the reasons, but could not alone explain the low repeatability. The facts that 95% confidence interval was narrower in DWI-EPI than that in DWI-TSE, and case 2 showed the highest repeatability even though the lesion was located in a region prone to distortion artifact (Fig. 1) support this interpretation. Identical distortion artifact deteriorates image quality but does not reduce voxel-wise ADC repeatability.

The data range between minimum and maximum narrowed with increasing NEX for ADC, but was not evident for T2 (Fig. 4). We interpret these findings as follows: DWI has such a wide data range (high data variability) and low repeatability that increasing NEX would narrow data range and increase repeatability. On the contrary, the data range of T2 is probably in plateau and related to the inhomogeneity of B0 and B1, so that increasing NEX does not alter data range.

5 Conclusions

The histogram metrics of voxel-wise ADC like minimum, 25%, median, 75%, maximum, and mean show high repeatability, but skewness and kurtosis do not. Radiologists should consider these characteristics when interpreting DWI and related metrics.

Acknowledgements

The authors would like to express special thanks to professor Mitsuru Mori for his statistical guidance. This work was supported in part by Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science (JSPS) (Grant Number 26461831). The part of this study was presented at the 44th annual meeting of the Japanese Society for Magnetic Resonance in Medicine.

Disclosure of Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

1. Hatakenaka M, Soeda H, Yabuuchi H, Matsuo Y, Kamitani T, Oda Y, Tsuneyoshi M, Honda H. Apparent diffusion coefficients of breast tumors: clinical application. Magn Reson Med Sci 2008; 7: 23-29.
2. Hatakenaka M, Shioyama Y, Nakamura K, Yabuuchi H, Matsuo Y, Sunami S, Kamitani T, Yoshiura T, Nakashima T, Nishikawa K, Honda H. Apparent diffusion coefficient calculated with relatively high b-values correlates with local failure of head and neck squamous cell carcinoma treated with radiotherapy. AJNR Am J Neuroradiol 2011; 32: 1904-1910.
3. Hatakenaka M, Nakamura K, Yabuuchi H, Shioyama Y, Matsuo Y, Kamitani T, Yonezawa M, Yoshiura T, Nakashima T, Mori M, Honda H. Apparent diffusion coefficient is a prognostic factor of head and neck squamous cell carcinoma treated with radiotherapy. Jpn J Radiol 2014; 32: 80-89.
4. Zolal A, Juratli TA, Linn J, Podlesek D, Sitoci Ficici KH, Kitzler HH, Schackert G, Sobottka SB, Rieger B, Krex D. Enhancing tumor apparent diffusion coefficient histogram skewness stratifies the postoperative survival in recurrent glioblastoma multiforme patients undergoing salvage surgery. J Neurooncol 2016; 127: 551-557.
5. Poussaint TY, Vajapeyam S, Ricci KI, Panigrahy A, Kocak M, Kun LE, Boyett JM, Pollack IF, Fouladi M. Apparent diffusion coefficient histogram metrics correlate with survival in diffuse intrinsic pontine glioma: a report from the Pediatric Brain Tumor Consortium. Neuro Oncol 2016; 18: 725-734.
6. Rodriguez Gutierrez D, Awwad A, Meijer L, Manita M, Jaspan T, Dineen RA, Grundy RG, Auer DP. Metrics and textural features of MRI diffusion to improve classification of pediatric posterior fossa tumors. AJNR Am J Neuroradiol 2014; 35: 1009-1015.
7. Nowosielski M, Recheis W, Goebel G, Güler O, Tinkhauser G, Kostron H, Schocke M, Gotwald T, Stockhammer G, Hutterer M. ADC histograms predict response to anti-angiogenic therapy in patients with recurrent high-grade glioma. Neuroradiology 2011; 53: 291-302.
8. Zhang YD, Wang Q, Wu CJ, Wang XN, Zhang J, Liu H, Liu XS, Shi HB. The histogram analysis of diffusion-weighted intravoxel incoherent motion (IVIM) imaging for differentiating the gleason grade of prostate cancer. Eur Radiol 2015; 25: 994-1004.
9. Wang Q, Li H, Yan X, Wu CJ, Liu XS, Shi HB, Zhang YD. Histogram analysis of diffusion kurtosis magnetic resonance imaging in differentiation of pathologic Gleason grade of prostate cancer. Urol Oncol 2015; 33: 337 e315-324. DOI: 10.1016/j.urolonc.2015.05.005

別刷請求先：小野寺　耕一
〒060-8543　札幌市中央区南 1 条西 16 丁目
札幌医科大学附属病院　臨床教育研究棟 3 階
放射線診断
TEL：011-611-2111（内線 35010）
FAX：011-633-6885
E-mail：koichi-onodera@sapmed.ac.jp

ADC ヒストグラムの指標に再現性はあるか？

小野寺耕一[1)]，山　直也[1)]，小野寺麻希[1)]，小山奈緒美[1)]，
喜友名由記[1)]，中西光広[2)]，畠中正光[1)]

[1)] 札幌医科大学医学部放射線診断学
[2)] 札幌医科大学附属病院放射線部

目的：本研究では臨床使用されている MRI において，見かけの拡散係数（ADC）を用いたヒストグラムの指標の再現性を検証した．

方法：2016 年 5 月から 7 月に当院にて頭部 MRI を施行した患者 12 名が本研究に参加し，すべての患者からインフォームドコンセントを取得した．同一ポジションで 2 回連続で撮像された echo planar imaging を用いた拡散強調画像（DWI-EPI）が得られた．b 値は 0，及び 1000 または 1500 s/mm^2 が使用され，3 軸直行の motion proving gradients（MPGs）を用いて，3 軸合成の拡散強調画像が生成された．関心領域（ROIs）が 1 回目の拡散強調画像の病変部位に置かれ，2 回目の拡散強調画像にも同じサイズや位置でコピー&ペーストされた．ボクセル毎の ADC 値は mono-exponential curve によるフィッティングを用いて計算された．1 回目と 2 回目の拡散強調画像の ADC 値はそれぞれ ADC-1st および ADC-2nd と定義された．それぞれの病変部位のボクセル毎の ADC 値の再現性を検証するため に，ADC-1st と ADC-2nd が Wilcoxon matched-pairs signed rank test と linear regression を用いて比較された．すべての病変で ADC ヒストグラムの指標の再現性を検証するために linear regression と Bland-Altman plot を使用し，ADC-1st と ADC-2nd におけるそれぞれの病変部位の minimal，25%，median，75%，maximum，mean，skewness，および kurtosis を検討した．

結果：ボクセル毎の ADC の再現性については，5 病変で ADC-1st と ADC-2nd に有意差を認めた．Linear regression では 5 病変で有意相関を認めなかった．ADC ヒストグラムの指標の再現性については，skewness と kurtosis を除いたすべての指標で linear regression における有意相関（$p<0.0001$）と Bland-Altman plot における高い再現性を認めた．

結 論：minimum，25%，median，75%，maximum，および mean の様なボクセル毎の ADC 値に基づくヒストグラムの指標は高い再現性を示したが，skewness と kurtosis は高い再現性を示さなかった．