Diagn Interv Radiol 2013; 19:433–437

Comparison of apparent diffusion coefficient values among different MRI platforms: a multicenter phantom study

Ali Sami Kıvrak, Yahya Paksoy, Cengiz Erol, Mustafa Koplay, Seda Özbek, Fatih Kara

PURPOSE
We aimed to compare apparent diffusion coefficient (ADC) values among magnetic resonance imaging (MRI) scanners from different vendors.

MATERIALS AND METHODS
We used a custom-made phantom solution consisting of distilled water, 0.9% NaCl, 25% NaCl, and shampoo for diffusion-weighted MRI (DW-MRI) examinations. DW-MRI was performed with similar sequence parameters using six different 1.5 Tesla MR scanners (scanners A–F). ADC maps were automatically constructed for all DW-MR images (b factors of 0 and 1000 s/mm^2). ADC measurements were performed using regions of interest and seven different software programs, including four different postprocessing workstations, two different picture archiving and communication systems, and operator console software for each MR scanner.

RESULTS
The ADC values generated by scanners A and F were higher and those of scanner B were lower than those generated by the other scanners ($P = 0.002$). The intravendor difference in the ADC values averaged from scanners D, E, and F was statistically significant ($P < 0.001$). The difference between the ADC values obtained by scanners C and E was not statistically different ($P = 0.15$).

CONCLUSION
ADC values may differ among different MRI systems used for DW-MRI. Thus, the MRI vendor should be considered when using DW-MRI in a clinical setting.

From the Departments of Radiology (A.S.K. ✉ *alisamikivrak@hotmail.com*, Y.P., C.E., M.K., S.Ö.) and Public Health (F.K.), Selçuk University School of Medicine, Konya, Turkey.

Received 30 January 2013; revision requested 27 February 2013; final revision received 14 May 2013; accepted 15 May 2013.

Published online 5 September 2013.
DOI 10.5152/dir.2013.13034

Diffusion-weighted magnetic resonance imaging (DW-MRI) depends on microscopic mobility that can be quantified using an apparent diffusion coefficient (ADC), or qualitatively assessed in biologic tissues. This mobility, classically called Brownian motion, includes molecular diffusion of water, which is the random thermal agitation of molecules and microcirculation of blood in capillary vessels. This motion causes phase dispersion of the spins, resulting in signal loss with the use of diffusion-sensitive sequences. This signal loss can be quantified using different b values to calculate the ADC, which depends largely on the presence of barriers to diffusion within the water microenvironment, namely, cell membranes and macromolecules (1–4).

Initially, the most important clinical application of DW-MRI was the detection and characterization of cerebral ischemia and tumors (5). With the development of echo-planar imaging and parallel imaging techniques, fast imaging times have become possible with MRI. DW-MRI is now widely used to image extracranial sites, including the liver, lung, and other organs, without the use of gadolinium chelates (2–4, 6, 7). It is mainly used to differentiate between benign and malignant lesions, monitor treatment response, and detect recurrent cancer on oncologic imaging with obtained ADC values (1). Many DW-MRI studies have reported that the ADC values of benign tumors are significantly higher than those of malignant tumors. This can probably be attributed to the combination of higher cellularity, tissue disorganization, and increased extracellular space of malignant masses (1, 2, 4). However, ADC values of some tumors, such as lung and liver carcinomas, may overlap. Under this condition, receiver operative characteristic (ROC) curve analyses are performed to determine cutoff ADC values to differentiate between benign and malignant lesions. Recently, cutoff points for ADC values have increasingly been suggested to differentiate between benign and malignant lesions in different organs (3, 4, 7, 8). However, reliable and reproducible cutoff values have not yet been confirmed. It is problematic that the ADC value varies with the scanning parameters of DW-MRI systems and vendor-specific issues such as field inhomogeneities, eddy currents, and sequence designs (9–11). Thus, serious problems and wrong conclusions may result when ADC is used to characterize tissue. Although similar scanning parameters are used for DW-MRI, ADC values vary, and standardization among different scanners has not been well established. In addition, analyses of DW-MRI data in multicenter trials should be performed at a single center using standardized validated software (1). However, it is unclear whether there has been consistency among the software used in published studies.

A multicenter study performed by Sasaki et al. (9) demonstrated that the intra- and intervendor ADC values can vary among different coil systems, scanners, vendors, and magnetic field strengths in healthy

volunteers. However, that study compared individual ADC values of the human brain. Nevertheless, the ability to differentiate between ADC values of different tissues is more important than to validate the use of DW-MRI as an imaging biomarker (1). Therefore, in multicenter trials using similar DW-MRI scanning methods, a comparison of precision and accuracy should be made using phantom solutions (hereafter, phantoms) to provide a basis for the pooling of data and sensitivity to motion effects not observed in phantoms. Improved anisotropic phantoms that reasonably mimic the cellular environment of living tissue are required to validate ADC measurements in multicenter clinical studies. However, it is difficult to develop a stable phantom because the cellular environment of living tissues cannot be easily mimicked (1, 12). For these reasons, we performed a multicenter phantom study and compared ADC values obtained from six different 1.5 Tesla (T) MR scanners by four different vendors with similar DW-MRI parameters using the same simple liquid isotropic phantoms. Our aim was to determine whether variation in ADC values exists among different MR scanners.

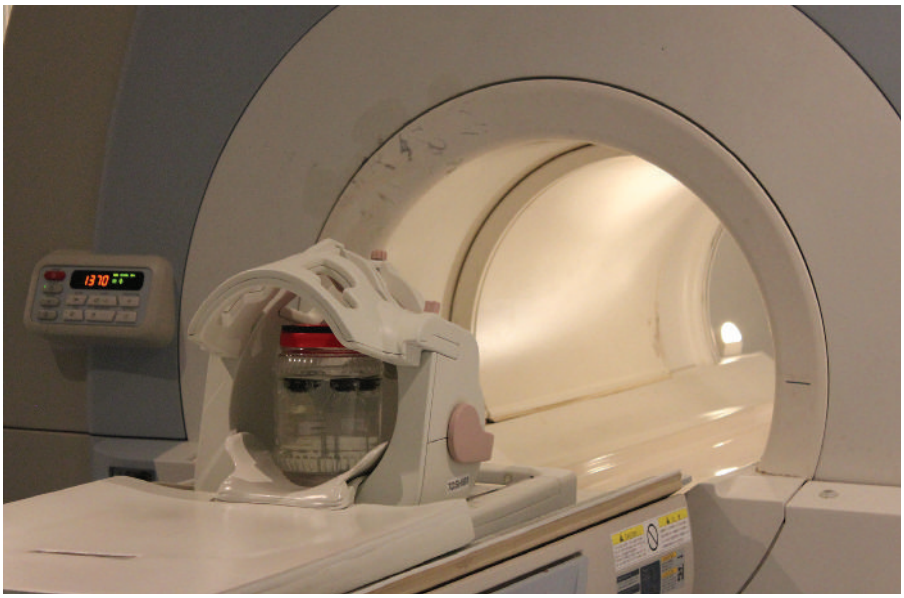

Figure 1. The phantom container containing four different fluids in 100 mL glass bottles was placed into the head coil in the MR scanner.

Materials and methods

Subjects

We used a homemade phantom for DW-MRI examination and compared ADC values that were obtained from six different MR scanners. As suggested in the literature, phantom materials used in multicenter clinical studies for DW-MRI should be inexpensive, easy to prepare, safe to transport, and stable over time such as water or a gel (1). Thus, simple liquid isotropic phantoms were used instead of anisotropic phantoms in the present study to provide a similar environment as that provided in a previous study (12). The phantom and methods used in the present study were as follows: four cylindrical glass bottle caps (100 mL volume, 5 cm diameter) containing distilled water, 0.9% NaCl solution, 25% NaCl solution, and shampoo were placed into a plastic phantom container containing tap water. Fluids, except shampoo, were prepared in the biochemistry laboratory. Shampoo was chosen because it is a moderately high-density fluid with a homogeneous structure. The phantom container was placed inside the MRI coil in the same direction and orientation; y-axis perpendicular in all of the MR scanners. Details of phantom positioning are shown in Fig. 1.

DW-MRI of the phantom was performed using six different 1.5 T MR scanners from four different vendors (Excelert Vantage, Toshiba Medical Systems, Tokyo, Japan [scanner A]; Signa, GE Healthcare, Milwaukee, Wisconsin, USA [scanner B]; Gyroscan Intera, Philips Healthcare, Best, the Netherlands [scanner C]; Avanto, Siemens Medical Systems, Erlangen, Germany [scanner D]; Symphony, Siemens Medical Systems [scanner E]; and Essenza, Siemens Medical Systems [scanner F]) at five different institutions. We used the same phantom for all hospitals and transferred them from one to another in the same day. The room temperature of the MRI platforms was 21°C. The phantom was allowed to sit at room temperature for 30 min before examination to standardize temperature. The phantom was examined just once on the same scanner. The acquired data were anonymized and collected in the Digital Imaging and Communications in Medicine (DICOM) format. ADC measurements were performed on seven different software and/or workstations, including four different postprocessing workstations (Vitrea 2, Vital Images, Minnetonka, Minnesota, USA; Virtual Place Advance Plus, Aze, Tokyo, Japan; Advantage Windows, GE Medical Systems; Leonardo, Siemens Medical Systems), two different picture archiving and communication systems (PACS) (Merge eFilm, Merge Technologies Inc., Milwaukee, Wisconsin, USA; Enlil PACS Eroğlu, Eskişehir, Turkey), and operator console software for each MR scanner. A circular region of interest (ROI) was set on every fluid of the ADC map of the phantom, and ADCs were measured and recorded. The ADC value was calculated for each pixel of the image, and then was displayed as a parametric map on all software. Intrascanner differences in the ADC values were not investigated using the different coil systems in the present study. Thus, sensitivity correction techniques were not applied to the MR images before ADC calculations.

Imaging protocol

DW-MRI examinations were performed using a 1.5 T MRI magnet and a multichannel head coil system on the six different platforms. All DW-MR images were obtained in the transverse plane with a multislice, single-shot, spin-echo echo-planar imaging DWI sequence using similar sequence parameters. DW-MR images were obtained at b factors of 0 and 1000 s/mm^2 for each section in the same sequence. Image acquisition schemes are summarized in Table 1. We used the automatic multiangle-projection shim and fat-suppression technique to

Table 1. Sequence parameters used for DW-MRI on different MR systems

Parameters	Excelert Vantage, Toshiba Medical Systems	Signa, GE Healthcare	Gyroscan Intera, Philips Healthcare	Avanto, Siemens Medical Systems	Symphony, Siemens Medical Systems	Essenza, Siemens Medical Systems
Field strength (T)	1.5	1.5	1.5	1.5	1.5	1.5
Coil (Head)	Multichannel 7 channel	Multichannel 4 channel	Multichannel 4 channel	Multichannel 8 channel	Multichannel 2 channel	Multichannel 6 channel
Slew rate (T/m/s)	130	120	150	125	125	100
Maximum gradient strength (mT/m)	30	33	30	33	30	30
Repetition time	6000	6000	6000	6000	6000	6000
Echo time[a]	100	130	100	102	118	111
Slice thickness (mm)	5	5	5	5	5	5
Field of view (mm)	230	230	230	230	230	230
Matrix size	128×128	128×128	128×128	128×128	128×128	128×128
b values (s/mm^2)	0, 1000	0, 1000	0, 1000	0, 1000	0,1000	0, 1000
Number of excitation	4	4	4	4	4	4

[a]Minimum echo time depends on the system.
mT/m, millitesla per meter; T, tesla.

reduce the artifacts in DW-MR images and applied an integrated phase correction during DW-MRI. We obtained 12–16 contiguous axial sections. ADC maps were automatically constructed for all DW-MR images (b factors of 0 and 1000 s/mm^2) using pixel-by-pixel calculations on the main MRI console of each device using software.

Data analysis

Analysis and ADC value measurements of DW-MRI data were obtained using the workstations and consoles to determine whether a significant difference exists between ADCs calculated by the different software programs. First, measurements were obtained on the main MRI consoles at each institution. Next, the ADC values of the fluids were calculated again on the images exported to the workstations, which was done to free up the main MRI console. Using a DICOM viewer, the same radiologist (A.S.K) measured the signal intensity through gray-scale ADC maps from each of the bottles into the phantom at a b=1000 s/mm^2 diffusion gradient value using a mouse-driven cursor to determine a circular ROI of 400 mm^2. Selecting a narrower or wider ROI would not affect ADC values, but ADC values were calculated using a similar ROI to mirror as closely as possible the measured region on all of the software. ROIs were placed at three different levels—upper, middle, and lower—without artifacts on the center of each bottle, and the mean ADC value was calculated as the average of three measurements. Forty-two mean ADC values were obtained for each phantom using seven different software programs and six different MR scanners.

Statistical analysis

Statistical analysis was performed using a commercially available software (Statistical Package for Social Sciences, version 15.0, SPSS Inc., Chicago, Illinois, USA). The Kruskal-Wallis test was used to contrast the ADC values of the four phantoms among different scanners. We determined whether significant differences existed among the seven median ADC values that were obtained for each phantom using the different software programs and the Kruskal-Wallis test. The Bonferroni-corrected Mann-Whitney U test was used to compare pairs of scanners using the median ADC values that were obtained for each phantom. Bonferroni correction test was used to identify significant differences among the different scanners and software programs (α=0.003). A difference with a P value less than 0.05 was deemed to be statistically significant for all analyses.

Results

All DW-MRI examinations were completed successfully. For all phantoms, the ADC maps were generated from the data obtained from each of the scanners used in the study (Fig. 2). Quantitative analysis demonstrated that the ADC calculation of one of the software programs (Vitrea 2, Vital Images) for scanners A and B was four times lower than that calculated by the other software programs, indicating a software vendor-based problem. Thus, the data of this software were not included in the statistical evaluation. There was no statistically significant difference among the other different software programs ($P > 0.05$). While the highest median ADC was found on scanner A, the lowest median ADC was found on scanner B for all fluids. The highest median ADC was found for distilled water (2425×10^{-6} mm^2/s; range, 2350–2427×10^{-6} mm^2/s). The lowest median ADC was found for 25% NaCl solution (1265×10^{-6} mm^2/s; range, 1265–1275×10^{-6} mm^2/s). The results showed that the fluids have different restricted diffusion consistency with density and amount of molecules. Quantitative analysis demonstrated that the ADC values for each fluid of the same phantom varied significantly among some scanners. Those generated by scanners A and F were higher and those of scanner B were lower than those generated by the other scanners; the differences were statistically significant (P = 0.002). The intravendor difference in the ADC values averaged from scanners D, E, and F was statisti-

cally significant ($P < 0.001$). The difference between the ADC values obtained by scanners C and E was not statistically different ($P = 0.15$) (Table 2).

We also calculated the differences between the ADC values of the phantoms obtained from three dedicated workstations: two PACS systems and the operator console of the six different vendor scanners. There were no significant differences among the ADC values of the software programs used.

Discussion

The purpose of this multicenter trial was to evaluate the repeatability and reproducibility of ADC measurements across the same and/or different vendor scanners. To ensure consistent and widespread application of quantitative ADC measurements, reproducibility and interimager variability should be known. Unfortunately, to date, no accepted standards in measurement or analysis methods have been established (1, 6). Hence, we directly compared the ADC values of different fluids using multiple scanners from different vendors with same field strengths and similar coil systems. Our results indicate that interimager standardization may not be possible because there was significant variation in ADC values among the different scanners.

A similar study performed previously by other researchers using healthy human subjects reached the same conclusion (9). That study assessed intra- and interimager ADC measurement variability of the normal brain using comparable diffusion protocols on different MR platforms (same and/or different vendor) in a group of volunteers and found that both measure varied. In addition, previous reports have indicated that the ADC values may vary within a given vendor and among different magnetic field strengths (1.5 and 3.0 T) (9, 13, 14). Such variation suggests the need for greater standardization of imaging parameters to minimize the measurement variability across platforms, allowing more meaningful comparison of results and facilitating multicenter studies (1). To the best of our knowledge, this is the first multicenter study to directly compare the ADC values between different MRI scanners using phantoms. The use of liquid phantoms instead of human subjects provided a comparison between the non individual ADC values in our study. We found a significant difference in the ADC values of all fluids among some imagers, although we applied nearly similar parameters to each system. A significant difference was detected between inter- and intravendor scanners, with the exception of scanners C and E. In addition, the median ADC values of scanner A were significantly higher than those of other scanners for all phantoms. However, a previous study that also used the scanner A used in the present study found that the ADC values were 15% lower than those of the other 1.5 T scanners tested (9). This suggests that the ADC value might significantly differ between scanners of

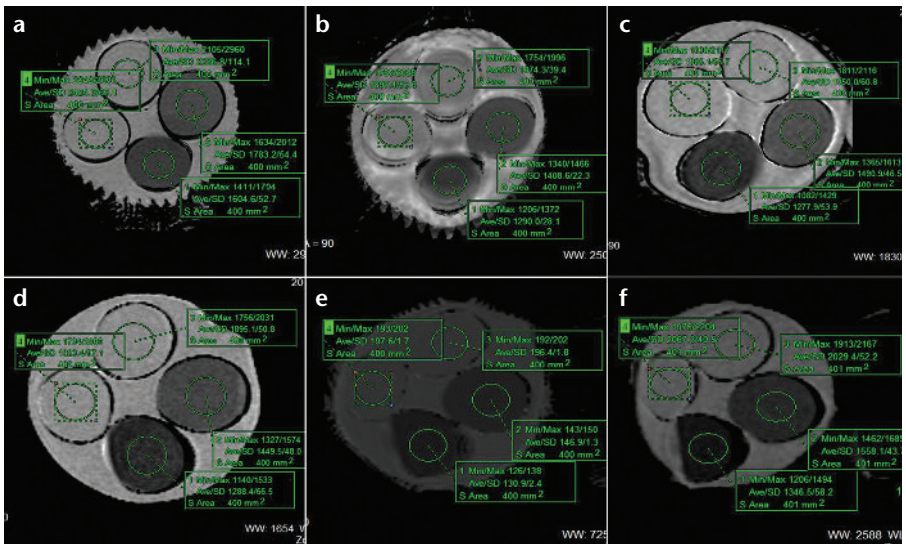

Figure 2. a–f. ADC maps from six different MRI scanners (**a**, Excelert Vantage, Toshiba Medical Systems; **b**, Signa, GE Healthcare; **c**, Gyroscan Intera, Philips Healthcare; **d**, Avanto, Siemens Medical Systems; **e**, Symphony, Siemens Medical Systems; **f**, Essenza, Siemens Medical Systems). ADC calculations were obtained from the workstation (Virtual Place Advance Plus, Aze).

Table 2. ADC values of the fluids on different scanners

	ADC value (x10^{-6} mm^2/s)			
Scanners	25% NaCl	Shampoo	0.9% NaCl	Distilled water
A	1579 (1555–1585)[a]	1785 (1727–1838)[a]	2377 (2356–2388)[a]	2425 (2350–2427)[a]
B	1265 (1265–1275)[a]	1398 (1382–1405)[a]	1848 (1835–1865)[a]	1885 (1860–1895)[a]
C	1284 (1275–1285)[b]	1465 (1460–1476)[b]	1950 (1945–1955)[b]	1955 (1945–1974)[b]
D	1275 (1270–1275)[c]	1465 (1443–1476)[c]	1898 (1890–1905)[c]	1915 (1915–1925)[c]
E	1290 (1280–1290)[b,c]	1450 (1450–1460)[b,c]	1940 (1930–1950)[b,c]	1950 (1940–1960)[b,c]
F	1335 (1335–1345)[a,c]	1560 (1552–1565)[a,c]	2066 (2055–2085)[a,c]	2070 (2065–2085)[a,c]

[a] $P = 0.002$ compared with the other vendor scanners.
[b] $P = 0.15$ compared with C and E scanners.
[c] $P < 0.001$ compared with same vendor scanners.
ADC, apparent diffusion coefficient; A, Excelert Vantage, Toshiba Medical Systems; B, Signa, GE Healthcare; C, Gyroscan Intera, Philips Healthcare; D, Avanto, Siemens Medical Systems; E, Symphony, Siemens Medical Systems; F, Essenza, Siemens Medical Systems.
Data are given as median (min–max).

even the same vendor due to a lack of standardization (due to scanner-specific protocols), complexities of data acquisition, differences among analysis methods, and/or as a result of the different software programs or PACS systems used.

DW-MRI data require unique approaches to quantify results. Standardized software would provide robust, standardized analyses (1). In the present study, we used dedicated workstations, main MRI consoles, and our PACS systems to calculate median ADC values and found no significant difference between median ADC values, as in a previous study (15). However, we found that the ADC values of one of the workstations were significantly different from the others for some scanners—we excluded these data from the statistical analysis. This suggests that data from some scanners may be incompatible with certain software, at least when calculating ADC values. Previous multicenter studies that reported variation among interimagers used only one workstation to calculate ADC values (9, 12). However, these studies did not test whether the workstation was compatible with different software programs. Therefore, our results are likely to be more reliable than those of previous studies in determining the difference among scanners.

Our study has some limitations. First, we used simple custom-made phantoms instead of volunteers or improved phantoms mimicking the cellular environment of living tissue. Thus, our phantoms might not be suitable for clinical practice. However, it is difficult to obtain ethical consent to perform repeated examinations on volunteers with multiple scanners. Nonetheless, using phantoms filled with beads may help us to understand the properties of diffusion of extracellular water but does not inform us about other water diffusion components such as vascular and intracellular contributions (1). Second, the number of materials needed to reach statistical accountability was relatively small because we used only four different fluids and six different scanners. However, we calculated the ADC values using six different software programs, and statistical analysis was carried out successfully. Third, distortion artifacts due to eddy currents at high b values may have influenced our DW-MRI measurements. However, statistically significant differences not occurring through mean ADC values were measured at three different levels without artifacts to minimize the effect of artifacts as much as possible. Fourth, susceptibility and magnetic field inhomogeneity-related artifacts could increase the variability of the calculated ADC. Even when the same MR system is used, DW-MRI studies have inherently lower signal-to-noise ratios and are susceptible to a range of artifacts (6). Finally, all ADC calculations were performed by the same author using similar ROIs. This may have affected median ADC values. However, our ADC calculations had less than 3% standard error.

DW-MRI is an attractive, noninvasive, qualitative, and quantitative technique that provides unique insight into tumor detection and characteristics, and there is growing evidence for its use in the assessment of patients with cancer. DW-MRI may be an effective early biomarker for treatment outcome for both antivascular drugs and therapies that induce tumor cell apoptosis (1, 2). To use it in this capacity in clinical practice, it is necessary to achieve consistency among all parties regarding standards for acquisition protocols, reproducibility, analysis methods, and software of the used workstation and/or PACS systems to ensure that quantitative ADC values have similar meanings across vendors and institutions.

In conclusion, there may be differences among ADC values obtained from different MRI systems used for DW-MRI. Thus, MRI system vendors should consider the clinical applications of DW-MRI to achieve accurate ADC measurements and a common language among radiologists. Standardization of not only protocols but also MRI system vendors and models for both image acquisition and data analysis across imaging platforms is important.

Conflict of interest disclosure

The authors declared no conflicts of interest.

References

1. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 2009; 11:102–125.
2. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. AJR Am J Roentgenol 2007; 188:1622–1635.
3. Matoba M, Tonami H, Kondou T, et al. Lung carcinoma: diffusion-weighted MR imaging--preliminary evaluation with apparent diffusion coefficient. Radiology 2007; 243:570–577.
4. Li Y, Chen Z, Wang J. Differential diagnosis between malignant and benign hepatic tumors using apparent diffusion coefficient on 1.5-T MR imaging: a meta analysis. Eur J Radiol 2012; 81:484–490.
5. Schaefer PW, Grant PE, Gonzalez RG. Diffusion-weighted MR imaging of the brain. Radiology 2000; 217:331–345.
6. Taouli B, Koh DM. Diffusion-weighted MR imaging of the liver. Radiology 2010; 254:47–66.
7. Turkbey B, Aras O, Karabulut N, et al. Diffusion-weighted MRI for detecting and monitoring cancer: a review of current applications in body imaging. Diagn Interv Radiol 2012; 18:46–59.
8. Kul S, Cansu A, Alhan E, Dinc H, Gunes G, Reis A. Contribution of diffusion-weighted imaging to dynamic contrast-enhanced MRI in the characterization of breast tumors. AJR Am J Roentgenol 2011; 196:210–217.
9. Sasaki M, Yamada K, Watanabe Y, et al. Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: a multivendor, multi-institutional comparison study. Radiology 2008; 249:624–630.
10. Dale BM, Braithwaite AC, Boll DT, Merkle EM. Field strength and diffusion encoding technique affect the apparent diffusion coefficient measurements in diffusion-weighted imaging of the abdomen. Invest Radiol 2008; 45:104–108.
11. Ogura A, Hayakawa K, Miyati T, Maeda F. Imaging parameter effects in apparent diffusion coefficient determination of magnetic resonance imaging. Eur J Radiol 2011; 77:185–188.
12. Chenevert TL, Galbán CJ, Ivancevic MK, et al. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. J Magn Reson Imaging 201; 34:983–987.
13. Bonekamp D, Nagae LM, Degaonkar M, et al. Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. Neuroimage 2007; 34:733–742.
14. Huisman TA, Loenneker T, Barta G, et al. Quantitative diffusion tensor MR imaging of the brain: field strength related variance of apparent diffusion coefficient (ADC) and fractional anisotropy (FA) scalars. Eur Radiol 2006; 16:1651–1658.
15. El Kady RM, Choudhary AK, Tappouni R. Accuracy of apparent diffusion coefficient value measurement on PACS workstation: a comparative analysis. AJR Am J Roentgenol 2011; 196:280–284.