**Phase shifts:** The process where protons lose coherence between each other due to magnetic field inhomogeneities; diffusion gradients have the sole purpose of enhancing the dephasing of spins thereby accelerating signal losses that are only recovered for spins that have remained static.

T_2-shine through: A term denoting visibility or mildly hyperintensity of tissues/structures related to the intrinsic T_2-weighting of DW images; usually stated to be present when mild hyper intensity is seen on very high b value images (800-1000 sec/mm^2) and on ADC maps.

Appendix 2. Roadmap for DW-MRI Development as Response Biomarker
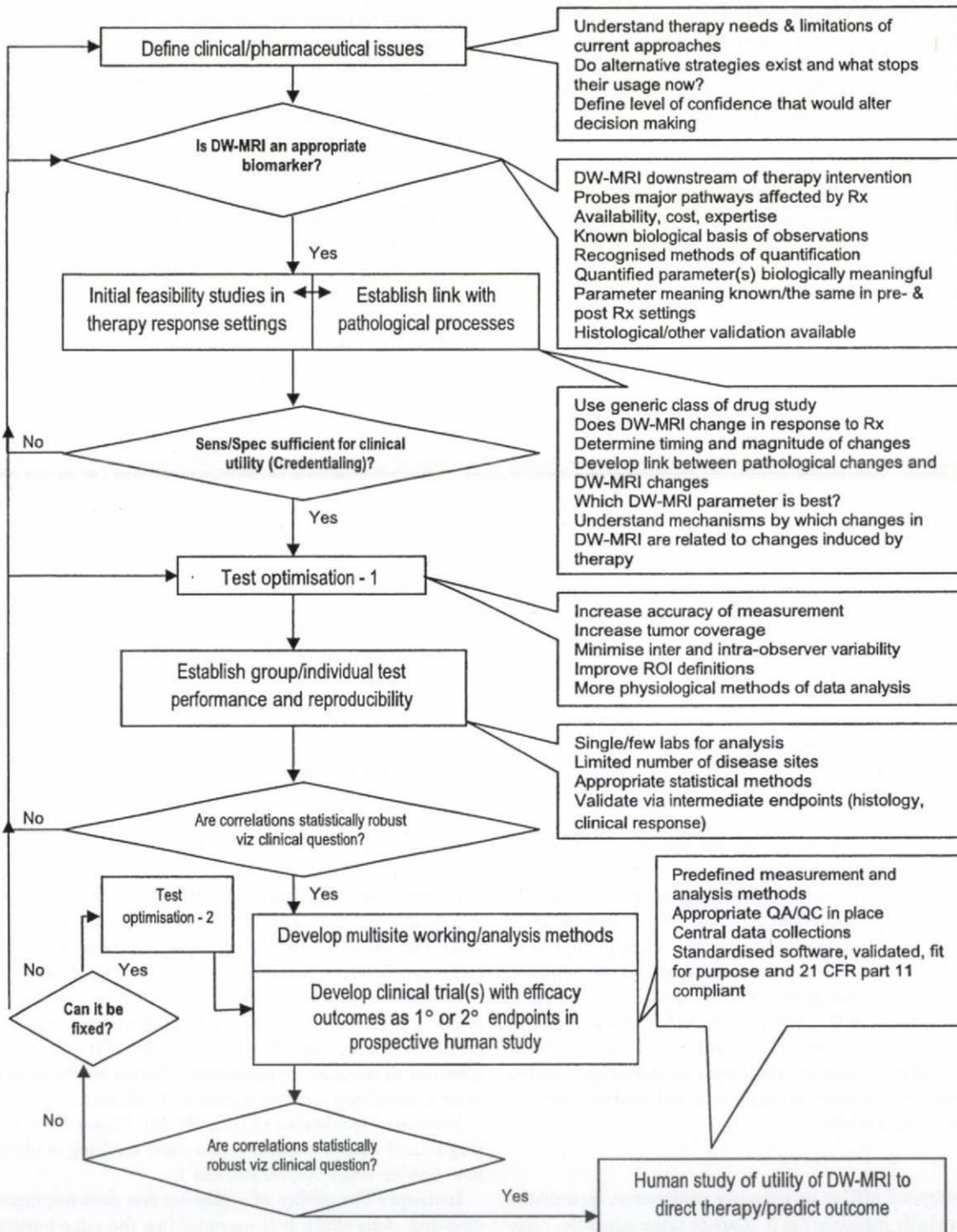
Appendix 3. Measurement Error Methodology

The International Standards Organization (ISO: 5725) definitions repeatability and reproducibility as follows:

1. *Repeatability* refers to test conditions that are as constant as possible, where the same operator using the same equipment within a "short time interval" obtains independent test results with the same method on identical items in the same laboratory.
2. *Reproducibility* refers to test conditions under which results are obtained with the same method on identical test items but in different laboratories with different operators using equipment.

Thus, repeatability refers to the ability of a measurement system to provide consistent readings on a given object. A major requirement is that external sources of error are controlled; in this case, only one observer performs the measurement and the spontaneous variability of the biologic parameter is assumed to change little during the brief period. Repeatability thus informs on equipment variation. In MRI, the repeated measurement of tissue relaxation properties without moving the patient over a short period is thus a repeatability study. It should be noted that repeatability is both a concept but also a specific statistical measure (see below).

Reproducibility, however, is the ability for multiple operators/experiments to achieve consistent results on identical objects. Thus, reproducibility measures the typical error between observers when each observes the same quantity. Reproducibility informs on appraiser/experimental variation. A DW-MRI measurement repeated on patients a few days apart is considered to be a reproducibility study. Even if the equipment and analysis technique used for the two measurements is identical, the experimental conditions are different because of the timing element.

Appropriate statistical parameters include the within-subject SD and coefficient of variance and interclass correlation coefficient. The repeatability statistic is a useful parameter for DW-MRI studies because it informs on whether changes in a particular patient are significant. Before statistical tests are applied, assumptions intrinsic to reproducibility analysis must be verified (normality of data and the nature of any relationship between measurement error and the magnitude of the parameters). Test-retest studies would need to be performed within a short period (the same day would do provided that patients are taken off the imaging couch for a period and a machine reboot takes place in a series of patients with the assumption that no significant changes are expected to occur over such a short period).

Measures of the spontaneous variability of parameter estimates

- Within subject coefficient of variation (wCV).
- Variance ratio (F): ratio of the between-patient variance and within-patient variance. A parameter with a larger variance in the patient population, but a small variance within individual patients (wCV) would have a higher variance ratio.

Measures of measurement error

- Intraclass correlation coefficient (ICC): a measure of the precision of parameter estimates (defining the reliability of the measurement method).
- Repeatability statistic: threshold value (absolute value or percent change) below which the absolute difference between

two measurements on the same patient ($n = 1$) is expected to lie for 95% of pairs of observations.

Statistical calculations also need to take into account that there may be more than one lesion measured per patient. Parameter values cannot be treated as independent samples because of the behavior of tumors within a particular patient shares to a degree, a biologic environment that is not common to other tumors. That is, some correlation between tumors in a patient is expected, and this clustering within patients needs to be taken into account in the methods of statistical analyses. The latter is likely to affect not only measurement error calculations but also the behavior of tumors owing to the effect of treatment. About the latter, it is likely that the behavior of tumors within a particular patient, due to treatment, is more likely to have a higher correlation than the variation between different patients.

Appendix 4. Diffusion Phantoms and QA Procedures

Quality control procedures for DW-MRI can characterize the performance of measurements obtained on clinical MRI systems and can assist in clinical studies performed in centers where there maybe limited specialist experience in diffusion imaging. Key features of QA protocols are as follows: 1) preparation of test objects (phantoms), 2) the systematic application of DW-MRI procedures, and 3) the parameters chosen to analyze the results of QA protocols must be helpful in identifying errors/deviations in the performance of MRI systems. There is debate on the best phantom material (options are outlined below), but whatever material is used, there is a pressing need for machine vendors to define clearly QA procedures for DW-MRI.

Test Object Manufacture

Ideally, test objects should be made of material with tissue-equivalent diffusions and MR signal properties (e.g., T_2-relaxation rate). Materials should be inexpensive, easy to prepare in a reproducible way, safe to transport, stable over time, and ideally nontoxic. However, it is recognized that multiple, complex phantoms to validate ADC measurements in multicenter clinical studies are difficult to make because the cellular environment of living tissues cannot be easily mimicked. Simple phantoms filled with gels/liquids with different diffusion coefficients may be acceptable initially.

A number of liquids have been suggested as substances for diffusion phantoms including the following:

1. Ice water is most dense, essentially 1.00 g/cm^3, at 4°C. This property makes ice water an attractive choice for a diffusion phantom because the measurement will not be temperature sensitive. Disadvantages include that fact that the free diffusion of ice water is not similar to that found in biologic tissues. The T1-relaxation rate of ice water is also long and requires measurements with long TRs. However, the safety and availability of this material are highly advantageous.
Water is a good choice of phantom for performing image quality evaluations because the geometric effects are not normally observed at the edge of phantoms, the problems associated with dielectric standing wave effects do not affect the geometry of the

sample, and the dielectric effects are independent of the effect of motion-probing gradients (MPGs).

2. Alkanes hydrocarbons such as decane, dodecane, and tetracane are thick, oily liquids of the paraffin series. These substances are harmful and may cause lung damage if swallowed. They should be used in original manufacturer's delivery bottles. Dodecane is nonhazardous for air, sea, and road freight and when used in a sealed state and it has diffusion values close to that of brain tissue [74].

3. Sucrose either used as an aqueous solution or used in combination with gels is a reliable test material whose properties can be readily adjusted to obtain tissue-like relaxations and diffusion properties [75,76]. Aqueous sucrose phantoms are easier to produce and are highly stable over time, although they require some form of antibacterial treatment during manufacture.

4. Polydimethylsiloxane is a high-viscosity, high–molecular weight polymer with a very low self-diffusion that is not detectable by DW-MRI; it also has a very low relative permittivity value. It is a useful material for evaluating a range of artifacts arising from DW-MRI but is not a useful material for determining ADC [77]. This is an excellent choice for phantom material because of the very low self-diffusion coefficient, resulting in no image attenuation resulting from the application of MPGs, its low dielectric constant eliminates standing wave problems; this is an advantage at increased static field strengths. However, it is also expensive.

   4.1. When considering phantom development, the white matter of normal brain has a highly reproducible ADC value (ranging from 60 to 105 $\times$ 10^{-2} mm^2/s [78,79] and might serve as a "traveling living phantom." Whereas the human brain has been proposed as a standard for calibration because of its stable temperature regulation, there are obvious limitations to using a human "phantom" in the context of multi-center trials.

Measurement Procedures

Whatever the final choice of diffusion phantom material used, the following measurement conditions should apply.

- A scanner-specific QA measurement protocol should be prepared in advance of the commencement of studies and should not be changed for the duration of the study.
- Imaging coils should be identical to those used in proposed clinical studies.
- Phantom position should be highly reproducible within the coils used for the QA procedure. This can be achieved either by the manufacture of positioning aids or with the use of bead-filled vacuum bags (often used to assure positioning in radiotherapy systems). Markers placed on phantoms can help to ensure that alignment and positioning are reproduced.
- The shape of the phantom should be either spherical or cylindrical.
- The temperature of the phantom and the scanner room should be recorded. Ideally, the phantoms should be left permanently in magnet rooms so that that temperature equilibrium is reached before scanning procedures are performed. Phantom ADC measurements need to be adjusted to take into account temperature because ADC increases by 2.5% per de-

gree Celsius [74]. The advantage of ice water as a test material if used as an alternative is that it is always measured at the same temperature and it eliminates temperature dependence from measurements.

- Clearly, the FOV, b values, MPG directions, imaging matrix, TR, TE, parallel imaging factor, number of slices, slice positions, and slice thickness must be held constant and be matched to clinical protocols. Data should be acquired without filters or image correction algorithms.
- All QA procedures should be performed with and without the use of fat suppression methods. The fat suppression method may affect image quality, and the effect of fat suppression method used should be evaluated.
- Because most clinical studies are performed with the acquisition of axial slices, the phantom setup should be designed to evaluate the QA DW-MRI measurements in this orientation.
- For all clinical DW-MRI protocols using multiple MPG directions, each MPG direction should be assessed independently (e.g., ADC_x, ADC_y, ADC_z) and ADC trace for the range of b values used in the clinical protocol, this could be restricted to two b values (b_0 and largest b value used). Variability should be established for each directional measurement. Apparent diffusion coefficient values should vary from less than 1% to 2% in each MPG direction for b values greater than b = 500 sec/mm^2.
- Apparent diffusion coefficient calculations may be affected by interactions of MPGs with the imaging gradients (a cross-term phenomenon) particularly at lower b values and on reduced FOVs.

Quality Assurance Analyses

- Analyses should be on ADC measurements obtained using methods described by Delakis et al. [75] about long-term fluctuations of ADC measurements and directional variations of ADC measurements.
- Diffusion decay curves should be evaluated to verify its monoexponential nature (b values should be sufficiently large >200 sec/mm^2 to reduce the effect of imaging gradient cross-terms on the actual b value). This is best achieved by performing a multiple averaged experiment, with TRs at least five times greater than the sample T1 to produce very high SNR diffusion decay curves. This should be evaluated in each direction that the MPGs are applied. Three-scan trace (tetrahedral) approaches, in general, do not enable users to evaluate the individual scan measurement data because only trace images are produced (it would be helpful if vendors provided an option to supply the directional images).
- In general, machine vendors do not provide a suitable software for evaluating the quality of the ADC fitting; this evaluation will require the use of dedicated software tools to perform evaluations (several software tools are freely available on the Internet). An ROI analysis can be performed using the vendor software, and the mean value per b value from a central ROI can be recorded and evaluated in a spreadsheet form using commonly available software (e.g., Matlab, Microsoft Excel, SPSS Inc. Sigmaplot, and many others). Significant data deviations from a monoexponential form or calculated ADC

estimates at variance with the vendor software results should be reported to vendors for remedial correction.

- Quality assurance measurements should be performed initially on a daily/weekly basis to establish the variability in the measurement data. Having established the variability for a given scanner, QA procedures should be performed routinely on a weekly basis and following any software/hardware changes.
- Records should be kept of QA analyses. Confidence limits should be established from the initial QA measurement data, and action procedures should be prescribed for QA measurements results that fall outside the confidence intervals [80].

Quantifying Artifacts in DW-MRI

Methods of quantifying artifacts in a systematic way are helpful in developing clinical protocols and maintaining image quality. Nondiffusing phantoms are well suited for this purpose. Apparent diffusion coefficient maps generated from measurements on nondiffusing test objects provide a direct indication of the source and magnitude of artifacts. Nyquist ghosts and distortions can be quantified [80], and protocols are optimized to reduce the magnitude of artifacts as indicated in the next section [77].

Using Phantoms to Improve DW-MR Image Quality

All DW-MRI measurements are influenced by artifacts and machine imperfections. These include b_0 inhomogeneity resulting from susceptibility variations within biologic or physical test objects/samples (this includes patients, volunteers, and nonhomogeneous structured phantoms). Chemical shift artifacts result from the presence of more than one chemical species or scalar coupling. Other artifacts are measurement-induced, for example, Nyquist ghosting and geometric distortions from residual MPG-induced eddy currents.

Before deciding on the measurement protocol for clinical studies, artifacts should be characterized, and their effects minimized using QA phantoms. There is extensive published literature on DTI informing on methods for improving DW-MR image quality [81,82]. Unfortunately, most successful DTI correction techniques are not easily applied in extracranial applications because of factors such as motion, patient-induced b_0 variations (including time variation), and chemical shift artifacts from fat.

Phantom measurements can be used to improve the quality of DW-MR images. One of the most challenging areas to tackle is geometrical distortion (which occurs over large FOVs typically used in extracranial imaging) resulting from residual eddy currents. Large homogeneous phantoms (>20 cm) are suitable for the evaluation of in-plane geometric distortions and have the advantage that they lack internal structures, which may confound the analysis [83]. Measurements performed on these phantoms are easy to evaluate using vendor measurement tools (distance and grid tools). Subtraction images (the DW image is rescaled by the ratio of mean values from the b_0 and DW images and then subtracted from the b_0 image) enable the degree of geometric distortion to be visualized easily.

Image distortion in clinical scanners can be improved in a number of ways:

1. By reducing the amplitude of the higher b values either directly or by using three-scan trace methods (these generally reduce the

amplitude of the individual gradients by a factor of 0.5-0.67 depending on vendor implementation).
2. Increasing the pixel bandwidth (image bandwidth, increased fatwater shift); this effectively reduces the length of the image readout and it may reduce image distortion and will reduce SNR. In practice, the bandwidth is normally increased until Nyquist ghosting becomes apparent in the images and is then subsequently reduced to an acceptable level.
3. Other choices include increasing the speed-up factor in the parallel imaging; factors greater than 2 can be problematic or unavailable depending on the scanner configuration.
4. Vendor-specific DW-MRI measurement methods exploiting the use of double refocused spin echo with balanced MPG (zero net MPG integral) are very effective in reducing eddy current–induced distortions; they do incur a minor increase in the minimum TE of the measurement [84].

Appendix 5. Sample 1.5-T Protocols

	Brain	Liver
Sequence type	SS-EPI	SS-EPI
FOV* (cm)	23 × 23	28-40
Matrix size* (x, y)	112 × 256	128 × 128; interpolate if possible
TR (msec)	Shortest = 4000	>2500
TE (msec)	89	Minimum
Fat suppression	SPIR	SPAIR/STIR
EPI factor	89	128
Parallel imaging factor	N/A	2
Signal averages	1	4-7
Section thickness (mm)/gap	5/0	5-7/0-1
Directions of MPGs	3 directions	3 directions
b factors (mm²/sec)	$b = 0, 1000$	Use three including $b = 0, 50,$ or 100, 500, or 1000
Pixel bandwidth (Hz)	1833	1446
Patient preparation	None	Empty stomach
Other comments	None	Free breathing or respiratory-triggered

Imaging protocols are for 1.5-T systems.
*Rectangular to fit body size and shape.
SPIR indicates spectral presaturation with inversion recovery; *WE*, water excitation.

	Pelvis	Whole body*
Sequence type	SS-EPI	SS-EPI
FOV† (cm)	26	28-40
Matrix size† (x, y)	160 × 256	128 × 128; interpolate if possible
TR (msec)	>3500	>2500
TE (msec)	Minimum	Minimum (<80)
Fat suppression	STIR	SPAIR/STIR
EPI factor	114	128
Parallel imaging	2	2
Signal averages	6	4-7
Section thickness (mm)/gap	6/1	5-7/0-1
Directions of MPGs	3	3
b factors (mm²/sec)	0, 100, 800	0, 50-100, 500-1000
Pixel bandwidth (Hz)	1000-1500	1000-1500
Patient preparation	Antiperistaltic – i.m. for longer action	Empty stomach
Other comments	Free breathing	Free breathing or respiratory-triggering

Imaging protocols are for 1.5-T systems.
*Koh DM, Takahara T, Imai Y, and Collins DJ (2007). Practical aspects of assessing tumors using clinical diffusion-weighted imaging in the body. *Magn Reson Med Sci* 6(4), 211–224.
†Rectangular to fit body size and shape.

Appendix 6. Acknowledgments and Workshop Participants List

Workshop Participants

The writing committee acknowledges workshop participants and thanks them for their ideas and thoughts on improving the quality of this White Paper

Anwar Padhani, MB, BS, FRCP, FRCR, Mount Vernon Hospital; Guoying Liu, PhD, National Cancer Institute; Peter L. Choyke, MD, National Cancer Institute; Thomas L. Chenevert, PhD, University of Michigan Health System; Brian D. Ross, PhD, University of Michigan Health System; Andy Dzik-Jurasz, MD, PhD, Novartis Pharmaceuticals Corporation; Dow-Mu Koh, Royal Marsden Hospital; Marc Van Cauteren, PhD, Philips Healthcare Asia Pacific; Bachir Taouli, MD, NYU Medical Center; Taro Takahara, MD, PhD, University Medical Center (UMC) Utrecht; A. Gregory Sorensen, MD, Harvard Medical School; Harriet C. Thoeny, University of Bern; Jeffry R. Alger, PhD, University of California, Los Angeles; David L. Buckley, PhD, The University of Manchester; Jean Brittain, GE Healthcare; Cecil Charles, Duke University; Haesun Choi, MD, M.D. Anderson Cancer Center; Laurence (Larry) Clarke, PhD, National Cancer Institute; David A. Clunie, RadPharm, Inc.; Alexandre Coimbra, Merck, Inc.; Patricia E. Cole, Etai Pharmaceuticals; David Collins, Royal Marsden Hospital; Amita Dave, PhD, Memorial Sloan-Kettering Cancer Center; Benjamin M. Ellingson, PhD, Marquette University, Medical College of Wisconsin; Lloyd Estkowski, GE Healthcare; Jeffrey L. Evelhoch, PhD, Amgen; John Froehlich, Guerbet Switzerland; Susan Galbraith, BSc, MB, BChir, PhD, Bristol-Myers Squibb; Robert J. Gillies, PhD, University of Arizona; Ingrid S. Gribbestad, Institutt for sirkulasjon og bildediagnostikk NTNU Norway; Masoom Haider, University of Toronto; Wendy Hayes, DO, Bristol-Myers Squibb Co.; Gwenael Herigault, PhD, Philips Healthcare; Patrice Hervo, GE Healthcare, France; Luna Hilaire, PhD, GE Healthcare Medical Diagnostics R&D; Derek Hill, IXICO Ltd.; Doug Hussey, Cedara Software, Inc.; Nola Hylton, PhD, University of California, San Francisco; Marko Ivancevic, PhD, Philips Medical Systems; Paula M. Jacobs, PhD, SAIC Frederick, National Cancer Institute; Yasushi Kaji, MD, Dokkyo Medical University, School of Medicine; Ihab R. Kamel, MD, PhD, Johns Hopkins Hospital; Berthold Kiefer, Siemens AG; Adrian Knowles, GE Healthcare; Jenny Kuhelj, Varian Medical Systems; Normand Laperriere, MD, FRCPC, Princess Margaret Hospital; Martin Leach, Royal Marsden Hospital; Cynthia Ménard, MD, FRCPC, Princess Margaret Hospital; Chuck Meyer, PhD, University of Michigan; Chrit Moonen, University Victor Segalen Bordeaux; Josephine Naish, University of Manchester; Sarah Nelson, UCSF – Mission Bay; Andrew N. Priest, Cambridge University Hospitals; Marwan Sati, Cedara Software, Inc.; Vahan Sharoyan, Synarc, Inc.; Girish Srinivasan; Choon Hua Thng, National Cancer Centre, Singapore; Paul S. Tofts, Brighton and Sussex Medical School; Nina Tunariu, Hammersmith Hospitals; J. C. Waterton, AstraZeneca; Souhil Zaim, MD, Synarc, Inc., University of Michigan Health System.

References

[1] Rudin M (2007). Imaging readouts as biomarkers or surrogate parameters for the assessment of therapeutic interventions. *Eur Radiol* 17 (10), 2441–2457.

[2] Johnston KC, Wagner DP, Wang XQ, Newman GC, Thijs V, Sen S, and Warach S (2007). Validation of an acute ischemic stroke model: does diffusion-weighted imaging lesion volume offer a clinically significant improvement in prediction of outcome?. *Stroke* 38 (6), 1820–1825.

[3] Chenevert TL, Brunberg JA, and Pipe JG (1990). Anisotropic diffusion in human white matter: demonstration with MR techniques *in vivo*. *Radiology* 177 (2), 401–405.

[4] Stejskal EO and Tanner JE (1965). Spin diffusion measurements: spin echoes in the presence of a time dependent field gradient. *J Chem Phys* 42, 288–292.

[5] Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, and Laval-Jeantet M (1988). Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 168, 497–505.

[6] Le Bihan D, Delannoy DJ, and Levin RL (1988). Temperature mapping with MR imaging of molecular diffusion: application to hyperthermia. *Radiology* 171, 853–857.

[7] Moseley ME, Cohen Y, Kucharczyk J, Mintorovitch J, Asgari HS, Wendland MF, Tsuruda J, and Norman D (1990). Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology* 176 (2), 439–445.

[8] Lee JH and Springer CS Jr (2003). Effects of equilibrium exchange on diffusion-weighted NMR signals: the diffusigraphic "shutter-speed". *Magn Reson Med* 49 (3), 450–458.

[9] Norris DG (2001). The effects of microscopic tissue parameters on the diffusion weighted magnetic resonance imaging experiment. *NMR Biomed* 14 (2), 77–93.

[10] Niendorf T, Dijkhuizen RM, Norris DG, van Lookeren Campagne M, and Nicolay K (1996). Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging. *Magn Reson Med* 36 (6), 847–857.

[11] Mulkern RV, Gudbjartsson H, Westin CF, Zengingonul HP, Gartner W, Guttmann CR, Robertson RL, Kyriakos W, Schwartz R, Holtzman D, et al. (1999). Multicomponent apparent diffusion coefficients in human brain. *NMR Biomed* 12 (1), 51–62.

[12] Bennett KM, Schmainda KM, Bennett RT, Rowe DB, Lu H, and Hyde JS (2003). Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn Reson Med* 50, 727–734.

[13] Thoeny HC and De Keyzer F (2007). Extracranial applications of diffusion-weighted magnetic resonance imaging. *Eur Radiol* 17 (6), 1385–1393.

[14] Patterson DM, Padhani AR, and Collins DJ (2008). Technology insight: water diffusion MRI—a potential new biomarker of response to cancer therapy. *Nat Clin Pract Oncol* 5 (4), 220–233.

[15] Hayashida Y, Yakushiji T, Awai K, Katahira K, Nakayama Y, Shimomura O, Kitajima M, Hirai T, Yamashita Y, and Mizuta H (2006). Monitoring therapeutic responses of primary bone tumors by diffusion-weighted image: initial results. *Eur Radiol* 16, 2637–2643.

[16] Guo Y, Cai YQ, Cai ZL, Gao YG, An NY, Ma L, Mahankali S, and Gao JH (2002). Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging. *J Magn Reson Imaging* 16, 172–178.

[17] Humphries PD, Sebire NJ, Siegel MJ, and Olsen OE (2007). Tumors in pediatric patients at diffusion-weighted MR imaging: apparent diffusion coefficient and tumor cellularity. *Radiology* 245 (3), 848–854.

[18] Squillaci E, Manenti G, Cova M, Di Roma M, Miano R, Palmieri G, and Simonetti G (2004). Correlation of diffusion-weighted MR imaging with cellularity of renal tumours. *Anticancer Res* 24 (6), 4175–4179.

[19] Manenti G, Di Roma M, Mancino S, Bartolucci DA, Palmieri G, Mastrangeli R, Miano R, Squillaci E, and Simonetti G (2008). Malignant renal neoplasms: correlation between ADC values and cellularity in diffusion weighted magnetic resonance imaging at 3 T. *Radiol Med* 113 (2), 199–213.

[20] Sugahara T, Korogi Y, Kochi M, Ikushima I, Shigematu Y, Hirai T, Okuda T, Liang L, Ge Y, Komohara Y, et al. (1999). Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. *J Magn Reson Imaging* 9, 53–60.

[21] Yoshikawa MI, Ohsumi S, Sugata S, Kataoka M, Takashima S, Mochizuki T, Ikura H, and Imai Y (2008). Relation between cancer cellularity and apparent diffusion coefficient values using diffusion-weighted magnetic resonance imaging in breast cancer. *Radiat Med* **26** (4), 222–226.

[22] Taouli B, Vilgrain V, Dumont E, Daire JL, Fan B, and Menu Y (2003). Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients. *Radiology* **226**, 71–78.

[23] Parikh T, Drew SJ, Lee VS, Wong S, Hecht EM, Babb JS, and Taouli B (2008). Focal liver lesion detection and characterization with diffusion-weighted MR imaging: comparison with standard breath-hold T2-weighted imaging. *Radiology* **246** (3), 812–822.

[24] Sinha S, Lucas-Quesada FA, Sinha U, DeBruhl N, and Bassett LW (2002). *In vivo* diffusion-weighted MRI of the breast: potential for lesion characterization. *J Magn Reson Imaging* **15**, 693–704.

[25] Woodhams R, Matsunaga K, Iwabuchi K, Kan S, Hata H, Kuranami M, Watanabe M, and Hayakawa K (2005). Diffusion-weighted imaging of malignant breast tumors: the usefulness of apparent diffusion coefficient (ADC) value and ADC map for the detection of malignant breast tumors and evaluation of cancer extension. *J Comput Assist Tomogr* **29**, 644–649.

[26] Sumi M, Sakihama N, Sumi T, Morikawa M, Uetani M, Kabasawa H, Shigeno K, Hayashi K, Takahashi H, and Nakamura T (2003). Discrimination of metastatic cervical lymph nodes with diffusion-weighted MR imaging in patients with head and neck cancer. *AJNR Am J Neuroradiol* **24**, 1627–1634.

[27] King AD, Ahuja AT, Yeung DKW, Fong DK, Lee YY, Lei KI, and Tse GM (2007). Malignant cervical lymphadenopathy: diagnostic accuracy of diffusion-weighted MR imaging. *Radiology* **245**, 806–813.

[28] Thoeny HC, De Keyzer F, Vandecaveye V, Chen F, Sun X, Bosmans H, Hermans R, Verbeken EK, Boesch C, Marchal G, et al. (2005). Effect of vascular targeting agent in rat tumor model: dynamic contrast-enhanced *versus* diffusion-weighted MR imaging. *Radiology* **237**, 492–499.

[29] Hamstra D, Rehemtulla A, and Ross BD (2007). Diffusion magnetic resonance imaging: a biomarker for treatment response in oncology. *J Clin Oncol* **25**, 4104–4109.

[30] Thoeny HC, De Keyzer F, Chen F, Ni Y, Landuyt W, Verbeken EK, Bosmans H, Marchal G, and Hermans R (2005). Diffusion-weighted MR imaging in monitoring the effect of a vascular targeting agent on rhabdomyosarcoma in rats. *Radiology* **234**, 756–764.

[31] Pickles MD, Gibbs P, Lowry M, and Turnbull LW (2006). Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer. *Magn Reson Imaging* **24**, 843–847.

[32] Yankeelov TE, Lepage M, Chakravarthy A, Broome E, Niermann K, Kelley M, Meszoely I, Mayer I, Herman C, and McManus K (2007). Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results. *Magn Reson Imaging* **25**, 1–13.

[33] Theilmann RJ, Borders R, Trouard TP, Xia G, Outwater E, Ranger-Moore J, Gillies RJ, and Stopeck A (2004). Changes in water mobility measured by diffusion MRI predict response of metastatic breast cancer to chemotherapy. *Neoplasia* **6**, 831–837.

[34] Kamel IR, Reyes DK, Liapi E, Bluemke DA, and Geschwind JF (2007). Functional MR imaging assessment of tumor response after 90Y microsphere treatment in patients with unresectable hepatocellular carcinoma. *J Vasc Intervent Radiol* **18**, 49–56.

[35] Uhl M, Saueressig U, van Buiren M, Kontny U, Niemeyer C, Köhler G, Ilyasov K, and Langer M (2006). Osteosarcoma: preliminary results of *in vivo* assessment of tumor necrosis after chemotherapy with diffusion- and perfusion-weighted magnetic resonance imaging. *Invest Radiol* **41**, 618–623.

[36] Mardor Y, Pfeffer R, Spiegelmann R, Roth Y, Maier SE, Nissim O, Berger R, Glicksman A, Baram J, Orenstein A, et al. (2003). Early detection of response to radiation therapy in patients with brain malignancies using conventional and high *b*-value diffusion-weighted magnetic resonance imaging. *J Clin Oncol* **21** (6), 1094–1100.

[37] Eichler AF, Batchelor TT, and Henson JW (2007). Diffusion and perfusion imaging in subacute neurotoxicity following high-dose intravenous methotrexate. *Neuro Oncol* **9** (3), 373–377.

[38] Zbaren P, Weidner S, and Thoeny HC (2008). Laryngeal and hypopharyngeal carcinomas after (chemo)radiotherapy: a diagnostic dilemma. *Curr Opin Otolaryngol Head Neck Surg* **16** (2), 147–153.

[39] Thoeny HC, De Keyzer F, Chen F, Vandecaveye V, Verbeken EK, Ahmed B, Sun X, Ni Y, Bosmans H, Hermans R, et al. (2005). Diffusion-weighted magnetic resonance imaging allows noninvasive *in vivo* monitoring of the effects

of combretastatin a-4 phosphate after repeated administration. *Neoplasia* **7**, 779–787.

[40] Dzik-Jurasz A, Domenig C, George M, Wolber J, Padhani A, Brown G, and Doran S (2002). Diffusion MRI for prediction of response of rectal cancer to chemoradiation. *Lancet* **360** (9329), 307–308.

[41] DeVries AF, Kremser C, Hein PA, Griebel J, Krezcy A, Ofner D, Pfeiffer KP, Lukas P, and Judmaier W (2003). Tumor microcirculation and diffusion predict therapy outcome for primary rectal carcinoma. *Int J Radiat Oncol Biol Phys* **56** (4), 958–965.

[42] Chenevert TL, Stegman LD, Taylor JM, Robertson PL, Greenberg HS, Rehemtulla A, and Ross BD (2000). Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors. *J Natl Cancer Inst* **92** (24), 2029–2036.

[43] Galons JP, Altbach MI, Paine-Murrieta GD, Taylor CW, and Gillies RJ (1999). Early increases in breast tumor xenograft water mobility in response to paclitaxel therapy detected by non-invasive diffusion magnetic resonance imaging. *Neoplasia* **1** (2), 113–117.

[44] Shen SH, Chiou YY, Wang JH, Yen MS, Lee RC, Lai CR, and Chang CY (2008). Diffusion-weighted single-shot echo-planar imaging with parallel technique in assessment of endometrial cancer. *AJR Am J Roentgenol* **190** (2), 481–488.

[45] Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S, and Van Cauteren M (2004). Diffusion weighted whole body imaging with background body signal suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display. *Radiat Med* **22** (4), 275–282.

[46] Kwee TC, Takahara T, Ochiai R, Nievelstein RA, and Luijten PR (2008). Diffusion-weighted whole-body imaging with background body signal suppression (DWIBS): features and potential applications in oncology. *Eur Radiol* **18** (9), 1937–1952.

[47] Schwamm LH, Koroshetz WJ, Sorensen AG, Wang B, Copen WA, Budzik R, Rordorf G, Buonanno FS, Schaefer PW, and Gonzalez RG (1998). Time course of lesion development in patients with acute stroke: serial diffusion- and hemodynamic-weighted magnetic resonance imaging. *Stroke* **29** (11), 2268–2276.

[48] Singer OC, Humpich MC, Fiehler J, Albers GW, Lansberg MG, Kastrup A, Rovira A, Liebeskind DS, Gass A, Rosso C, et al. (2008). Risk for symptomatic intracerebral hemorrhage after thrombolysis assessed by diffusion-weighted magnetic resonance imaging. *Ann Neurol* **63** (1), 52–60.

[49] Reddy JS, Mishra AM, Behari S, Husain M, Gupta V, Rastogi M, and Gupta RK (2006). The role of diffusion-weighted imaging in the differential diagnosis of intracranial cystic mass lesions: a report of 147 lesions. *Surg Neurol* **66** (3), 246–250; discussion 250–251.

[50] Schroeder PC, Post MJ, Oschatz E, Stadler A, Bruce-Gregorios J, and Thurnher MM (2006). Analysis of the utility of diffusion-weighted MRI and apparent diffusion coefficient values in distinguishing central nervous system toxoplasmosis from lymphoma. *Neuroradiology* **48** (10), 715–720.

[51] Engelhorn T, Hufnagel A, Weise J, Baehr M, and Doerfler A (2007). Monitoring of acute generalized status epilepticus using multilocal diffusion MR imaging: early prediction of regional neuronal damage. *AJNR Am J Neuroradiol* **28** (2), 321–327.

[52] Dechambre S, Duprez T, Lecouvet F, Raftopoulos C, and Gosnard G (1999). Diffusion-weighted MRI postoperative assessment of an epidermoid tumour in the cerebellopontine angle. *Neuroradiology* **41** (11), 829–831.

[53] Kono K, Inoue Y, Nakayama K, Shakudo M, Morino M, Ohata K, Wakasa K, and Yamada R (2001). The role of diffusion-weighted imaging in patients with brain tumors. *AJNR Am J Neuroradiol* **22** (6), 1081–1088.

[54] Sijens PE, Heesters MA, Enting RH, van der Graaf WT, Potze JH, Irwan R, Meiners LC, and Oudkerk M (2007). Diffusion tensor imaging and chemical shift imaging assessment of heterogeneity in low grade glioma under temozolomide chemotherapy. *Cancer Invest* **25** (8), 706–710.

[55] Hamstra DA, Chenevert TL, Moffat BA, Johnson TD, Meyer CR, Mukherji SK, Quint DJ, Gebarski SS, Fan X, Tsien CI, et al. (2005). Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma. *Proc Natl Acad Sci USA* **102** (46), 16759–16764.

[56] Mori S and van Zijl PC (2002). Fiber tracking: principles and strategies—a technical review. *NMR Biomed* **15** (7–8), 468–480.

[57] Goebell E, Paustenbach S, Vaeterlein O, Ding XQ, Heese O, Fiehler J, Kucinski T, Hagel C, Westphal M, and Zeumer H (2006). Low-grade and anaplastic gliomas: differences in architecture evaluated with diffusion-tensor MR imaging. *Radiology* **239** (1), 217–222.

[58] Stadlbauer A, Ganslandt O, Buslei R, Hammen T, Gruber S, Moser E, Buchfelder M, Salomonowitz E, and Nimsky C (2006). Gliomas: histopathologic evaluation

of changes in directionality and magnitude of water diffusion at diffusion-tensor MR imaging. *Radiology* **240** (3), 803–810.

[59] Rustin GJ, Nelstrop AE, Bentzen SM, Bond SJ, and McClean P (2000). Selection of active drugs for ovarian cancer based on CA-125 and standard response rates in phase II trials. *J Clin Oncol* **18** (8), 1733–1739.

[60] Bammer R (2003). Basic principles of diffusion-weighted imaging. *Eur J Radiol* **45** (3), 169–184.

[61] Koh DM, Brown G, Riddell AM, Scurr E, Collins DJ, Allen SD, Chau I, Cunningham D, deSouza NM, Leach MO, et al. (2008). Detection of colorectal hepatic metastases using MnDPDP MR imaging and diffusion-weighted imaging (DWI) alone and in combination. *Eur Radiol* **18** (5), 903–910.

[62] Nasu K, Kuroki Y, Sekiguchi R, and Nawano S (2006). The effect of simultaneous use of respiratory triggering in diffusion-weighted imaging of the liver. *Magn Reson Med Sci* **5** (3), 129–136.

[63] Koh DM, Scurr E, Collins DJ, Pirgon A, Kanber B, Karanjia N, Brown G, Leach MO, and Husband JE (2006). Colorectal hepatic metastases: quantitative measurements using single-shot echo-planar diffusion-weighted MR imaging. *Eur Radiol* **16** (9), 1898–1905.

[64] Yoshikawa T, Kawamitsu H, Mitchell DG, Ohno Y, Ku Y, Seo Y, Fujii M, and Sugimura K (2006). ADC measurement of abdominal organs and lesions using parallel imaging technique. *AJR Am J Roentgenol* **187** (6), 1521–1530.

[65] Yamada I, Aung W, Himeno Y, Nakagawa T, and Shibuya H (1999). Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. *Radiology* **210**, 617–623.

[66] Ichikawa T, Haradome H, Hachiya J, Nitatori T, and Araki T (1998). Diffusion-weighted MR imaging with a single-shot echoplanar sequence: detection and characterization of focal hepatic lesions. *AJR Am J Roentgenol* **170** (2), 397–402.

[67] Ichikawa T, Haradome H, Hachiya J, Nitatori T, and Araki T (1999). Diffusion-weighted MR imaging with single-shot echo-planar imaging in the upper abdomen: preliminary clinical experience in 61 patients. *Abdom Imaging* **24** (5), 456–461.

[68] Ichikawa T, Erturk SM, Motosugi U, Sou H, Iino H, Araki T, and Fujii H (2006). High–*B*-value diffusion-weighted MRI in colorectal cancer. *AJR Am J Roentgenol* **187** (1), 181–184.

[69] Naganawa S, Sato C, Kumada H, Ishigaki T, Miura S, and Takizawa O (2005). Apparent diffusion coefficient in cervical cancer of the uterus: comparison with the normal uterine cervix. *Eur Radiol* **15** (1), 71–78.

[70] Westin CF, Maier SE, Mamata H, Nabavi A, Jolesz FA, and Kikinis R (2002). Processing and visualization for diffusion tensor MRI. *Med Image Anal* **6** (2), 93–108.

[71] Tozer DJ, Davies GR, Altmann DR, Miller DH, and Tofts PS (2006). Principal component and linear discriminant analysis of T1 histograms of white and grey matter in multiple sclerosis. *Magn Reson Imaging* **24** (6), 793–800.

[72] Moffat BA, Chenevert TL, and Lawrence TS (2005). Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. *Proc Natl Acad Sci USA* **102**, 5524–5529.

[73] Lee S, Choi M, Kim H, and Park FC (2007). Geometric direct search algorithms for image registration. *IEEE Trans Image Process* **16** (9), 2215–2224.

[74] Tofts PS, Lloyd D, Clark CA, Barker GJ, Parker GJ, McConville P, Baldock C, and Pope JM (2000). Test liquids for quantitative MRI measurements of self-diffusion coefficient *in vivo*. *Magn Reson Med* **43** (3), 368–374.

[75] Delakis I, Moore EM, Leach MO, and De Wilde JP (2004). Developing a quality control protocol for diffusion imaging on a clinical MRI system. *Phys Med Biol* **49** (8), 1409–1422.

[76] Laubach HJ, Jakob PM, Loevblad KO, Baird AE, Bovo MP, Edelman RR, and Warach S (1998). A phantom for diffusion-weighted imaging of acute stroke. *J Magn Reson Imaging* **8** (6), 1349–1354.

[77] Price WS, Barzykin AV, Hayamizu K, and Tachiya M (1998). A model for diffusive transport through a spherical interface probed by pulsed-field gradient NMR. *Biophys J* **74** (5), 2259–2271.

[78] Pierpaoli C, Jezzard P, Basser PJ, Barnett A, and Di Chiro G (1996). Diffusion tensor MR imaging of the human brain. *Radiology* **201** (3), 637–648.

[79] Sener RN (2001). Diffusion MRI: apparent diffusion coefficient (ADC) values in the normal brain and a classification of brain disorders based on ADC values. *Comput Med Imaging Graph* **25** (4), 299–326.

[80] Simmons A, Moore E, and Williams SC (1999). Quality control for functional magnetic resonance imaging using automated data analysis and Shewhart charting. *Magn Reson Med* **41** (6), 1274–1278.

[81] Horsfield MA (1999). Mapping eddy current induced fields for the correction of diffusion-weighted echo planar images. *Magn Reson Imaging* **17** (9), 1335–1345.

[82] Jezzard P, Barnett AS, and Pierpaoli C (1998). Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. *Magn Reson Med* **39** (5), 801–812.

[83] Mattila S, Renvall V, Hiltunen J, Kirven D, Sepponen R, Hari R, and Tarkiainen A (2007). Phantom-based evaluation of geometric distortions in functional magnetic resonance and diffusion tensor imaging. *Magn Reson Med* **57** (4), 754–763.

[84] Reese TG, Heid O, Weisskoff RM, and Wedeen VJ (2003). Reduction of eddy-current–induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn Reson Med* **49** (1), 177–182.

Original Research _____

Comparison and Reproducibility of ADC Measurements in Breathhold, Respiratory Triggered, and Free-Breathing Diffusion-Weighted MR Imaging of the Liver

Thomas C. Kwee, MD,[1]* Taro Takahara, MD, PhD,[1] Dow-Mu Koh, MD, PhD,[2,3] Rutger A.J. Nievelstein, MD, PhD,[1] and Peter R. Luijten, PhD[1]

Purpose: To compare and determine the reproducibility of apparent diffusion coefficient (ADC) measurements of the normal liver parenchyma in breathhold, respiratory triggered, and free-breathing diffusion-weighted magnetic resonance imaging (DWI).

Materials and Methods: Eleven healthy volunteers underwent three series of DWI. Each DWI series consisted of one breathhold, one respiratory triggered, and two free-breathing (thick and thin slice acquisition) scans of the liver, at b-values of 0 and 500 s/mm^2. ADCs of the liver parenchyma were compared by using nonparametric tests. Reproducibility was assessed by the Bland–Altman method.

Results: Mean ADCs (in 10^{-3} mm^2/sec) in respiratory triggered DWI (2.07–2.27) were significantly higher than mean ADCs in breathhold DWI (1.57–1.62), thick slice free-breathing DWI (1.62–1.65), and thin slice free-breathing DWI (1.57–1.66) ($P < 0.005$). Ranges of mean difference in ADC measurement ± limits of agreement between two scans were −0.02–0.05 ± 0.16–0.24 in breathhold DWI, −0.14–0.20 ± 0.59–0.60 in respiratory triggered DWI, −0.03–0.03 ± 0.20–0.29 in thick slice free-breathing DWI, and −0.01–0.09 ± 0.21–0.29 in thin slice free-breathing DWI.

Conclusion: ADC measurements of the normal liver parenchyma in respiratory triggered DWI are significantly higher and less reproducible than in breathhold and free-breathing DWI.

DIFFUSION-WEIGHTED (DW) MAGNETIC RESONANCE (MR) imaging with apparent diffusion coefficient (ADC) measurements is an emerging technique for the detection and characterization of liver lesions (1–6). Three approaches can be used for DW MR imaging (DWI) of the liver: breathhold, respiratory triggered, and free-breathing scanning (2). Breathhold scanning requires only a relatively short examination time, but signal-to-noise ratio (SNR) may be compromised (2). Respiratory triggered scanning allows achieving higher SNR; however, at the expense of prolonged examination time (2). The recently developed concept of DWI under free breathing, also known as diffusion-weighted whole-body imaging with background body signal suppression (DWIBS) (7), is an emerging technique for whole-body oncological imaging (including the liver), and allows thin slice acquisition and multiple slice excitations (which is impossible in breathhold DWI), without seriously prolonging image acquisition time (as would happen in respiratory triggered DWI). Free-breathing DWI of the liver is relatively time-efficient and offers good SNR, although respiratory triggered scanning usually provides sharper anatomical details (2).

ADC measurements in breathhold, respiratory triggered, or free-breathing DWI can be helpful in the characterization of liver lesions (1–6), provided that they are accurate and reproducible. However, to our knowledge, ADC measurements of the liver in breathhold, respiratory triggered, and free-breathing DWI have never been compared before. Furthermore, it is still unknown whether ADC measurements of the liver are reproducible in each of the three DWI sequences.

The aims of this study were therefore to compare and to assess the reproducibility of ADC measurements of

[1]Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands.
[2]Cancer Research UK Clinical Magnetic Resonance Research Group, Institute of Cancer Research, Sutton, United Kingdom.
[3]Academic Department of Radiology, Royal Marsden Hospital, Sutton, United Kingdom.
*Address reprint requests to: T.C.K., University Medical Center Utrecht, Department of Radiology, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands. E-mail: thomaskwee@gmail.com
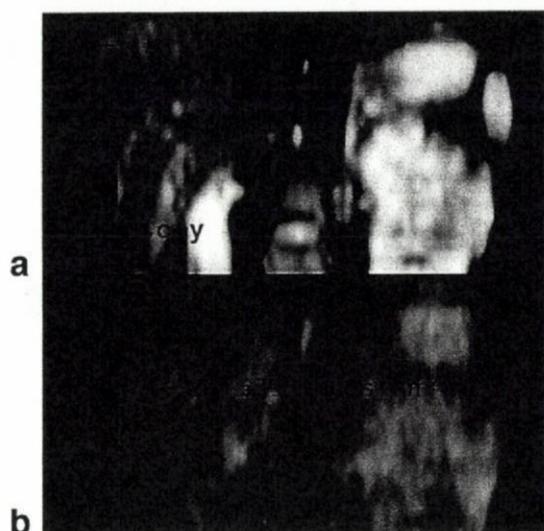
Figure 1. Example of ADC measurement in a coronally reformatted image. **a:** A rectangular region of interest (ROI), measuring 100 mm × 10 mm, is placed in the homogeneous hepatic parenchyma, at the lateral side of the right hepatic lobe, in the image with a b-value of 0 s/mm². **b:** The ROI is copied and pasted onto the ADC map, and ADC is automatically calculated.

the normal liver parenchyma in breathhold, respiratory triggered, and free-breathing DWI.

MATERIALS AND METHODS

Study Participants

This study was approved by the local Institutional Review Board and written informed consent was obtained from all participants. Eleven healthy adult volunteers (four men and seven women; mean age, 21.5 years; age range, 18–26 years) with no prior history or findings related to liver disease at the time of the study prospectively underwent DWI of the liver. Exclusion criteria were general contraindications to MR imaging, such as implanted pacemaker and claustrophobia.

MR Imaging

All volunteers were examined with a 1.5T MR scanner (Achieva, Philips Medical Systems, Best, The Netherlands) using a 16-element phased array surface coil (SENSE XL Torso coil, Philips Medical Systems). In each volunteer one (single, mid-expiratory) breathhold (slice thickness/gap: 7/1 mm), one (end-expiratory) respiratory triggered (slice thickness/gap: 7/1 mm), and two free-breathing (slice thickness/gap: 7/1 mm, and 5/0 mm) axial DW single-shot spin-echo echo-planar images of the liver were obtained, and this series was repeated two times. Thus, three series of one breathhold, one respiratory triggered, and two free-breathing DWI scans were obtained in each volunteer, resulting in a total of 12 scans in each volunteer. Applied sequence parameters for breathhold and respiratory triggered DWI were as follows: (repetition time [TR]/) echo time of (2247/) 70 msec, parallel imaging factor of 2, halfscan factor of 0.6, echo-planar imaging factor of 45, bandwidth of 1478.6 Hz, image acquisition in the axial plane, number of slices of 20, field of view (FOV) of 360 × 292 mm², rectangular FOV of 81.1%, matrix of 256, scan percentage of 80.4, actual voxel size of $2.8 \times 3.5 \times 7.0$ mm³, calculated voxel size of $1.4 \times 1.4 \times 1.7$ mm³, chemical shift selective fat suppression, 2 signals averaged. Thick slice free-breathing DWI (slice thickness/gap: 7/1 mm) was performed with a TR of 2248 msec and 5 signals averaged, and thin slice free-breathing DWI (slice thickness/gap: 5/0 mm) was performed with a TR of 3596 msec and 4 signals averaged, but other sequence parameters were identical to those of breathhold and respiratory triggered DWI. In each of the 12 DWI scans, motion probing gradients (MPGs) with two b-values (0 and 500 s/mm²) were applied in three directions: phase encoding (P), frequency encoding (M), and slice select (S). In each of the 12 DWI scans, five images were obtained at each level of the liver: one image at a b-value of 0 s/mm², P, M, and S images at a b-value of 500 s/mm², and one index isotropic (I) image at a b-value of 500 s/mm². Nominal scan times of breathhold DWI, respiratory triggered DWI, thick slice free-breathing DWI (slice thickness/gap: 7/1 mm), and thin slice free-breathing DWI (slice thickness/gap: 5/0 mm)

Table 1
Comparison of ADCs (in 10-3 mm2/s) of the Normal Liver Parenchyma

Sequence	Series	P	M	S	P value
BH	1	1.66 ± 0.19	1.61 ± 0.19	1.61 ± 0.21	0.779
	2	1.62 ± 0.16	1.65 ± 0.23	1.65 ± 0.23	0.993
	3	1.60 ± 0.14	1.54 ± 0.16	1.52 ± 0.21	0.664
RT	1	2.09 ± 0.68	2.26 ± 0.84	2.04 ± 0.44	0.993
	2	1.93 ± 0.23	2.77 ± 1.08	2.02 ± 0.38	0.069
	3	1.87 ± 0.27	2.29 ± 0.73	1.94 ± 0.47	0.429
FB7/1	1	1.71 ± 0.25	1.72 ± 0.18	1.63 ± 0.28	0.190
	2	1.59 ± 0.10	1.68 ± 0.16	1.61 ± 0.19	0.342
	3	1.64 ± 0.08	1.67 ± 0.18	1.61 ± 0.32	0.381
FB5/0	1	1.66 ± 0.19	1.61 ± 0.19	1.61 ± 0.21	0.686
	2	1.62 ± 0.16	1.65 ± 0.23	1.65 ± 0.23	0.139
	3	1.60 ± 0.14	1.54 ± 0.16	1.52 ± 0.21	0.092

Phase encoding (P), frequency encoding (M), and slice select (S) directions in coronally reformatted breathhold (BH), respiratory triggered (RT), free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and free breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0).

Table 2
Overall Comparison of Isotropic ADCs (in 10-3 mm2/s) of the Normal Liver Parenchyma

Series	BH	RT	FB 7/1	FB 5/0	P value
1	1.60 ± 0.14	2.13 ± 0.33	1.65 ± 0.09	1.64 ± 0.13	0.000
2	1.62 ± 0.18	2.27 ± 0.47	1.62 ± 0.16	1.66 ± 0.11	0.000
3	1.57 ± 0.15	2.07 ± 0.43	1.65 ± 0.17	1.57 ± 0.19	0.003

Coronally reformatted breathhold DWI (BH), respiratory triggered DWI (RT), free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and free breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0).

were 22.5 seconds, 24 seconds, 51.7 seconds, and 64.7 seconds, respectively. For respiratory triggered DWI, a commercially available air-filled pressure sensor (capable of measuring respiratory-induced pressure changes) was fixed to the hypochondrial region by an elastic belt. The respiratory cycle was automatically monitored by this device. Prior to scanning, volunteers were asked to perform normal, regular breathing, which was checked by the operator on the monitoring screen. Actual scan time of respiratory triggered DWI was manually recorded. Obtained axial images were coronally reformatted, with a slice thickness of 8 mm. ADC maps were created of coronally reformatted and axial images, using the b-values of 0 s/mm² and 500 s/mm².

Image Analysis

MR images were transferred to a workstation (ViewForum; Philips Medical Systems) and reviewed by consensus of a researcher and a radiologist with 14 years of experience in MR imaging. Only the right hepatic lobe was evaluated, because no cardiac gating was employed. Image quality of some axial slices can be poor;

selection bias is introduced when only (locations on) axial slices of good image quality are selected for analysis. Therefore, ADC measurements were initially only done on coronally reformatted images as follows: rectangular regions of interest (ROIs), measuring 100 × 10 mm, were created on all coronally reformatted breathhold, respiratory triggered, and free-breathing images at a b-value of 0 s/mm². Each ROI was placed in the homogeneous liver parenchyma (visible vascular and biliary structures were excluded), at the lateral side of the right hepatic lobe (with at margin of at least 0.5 mm from the lateral border of the liver), at the slice through the right portal vein. This approach was chosen because large vessels would inevitably have been included if the ROIs were placed in the medial part of the liver. Subsequently, all ROIs were copied and pasted onto the P, M, S, and I ADC maps, and ADCs were automatically calculated (Fig. 1).

Despite its previously mentioned disadvantage, ADC measurements were also done on axial images as follows: three circular ROIs, each measuring 20 mm², were created on all axial breathhold, respiratory trig-
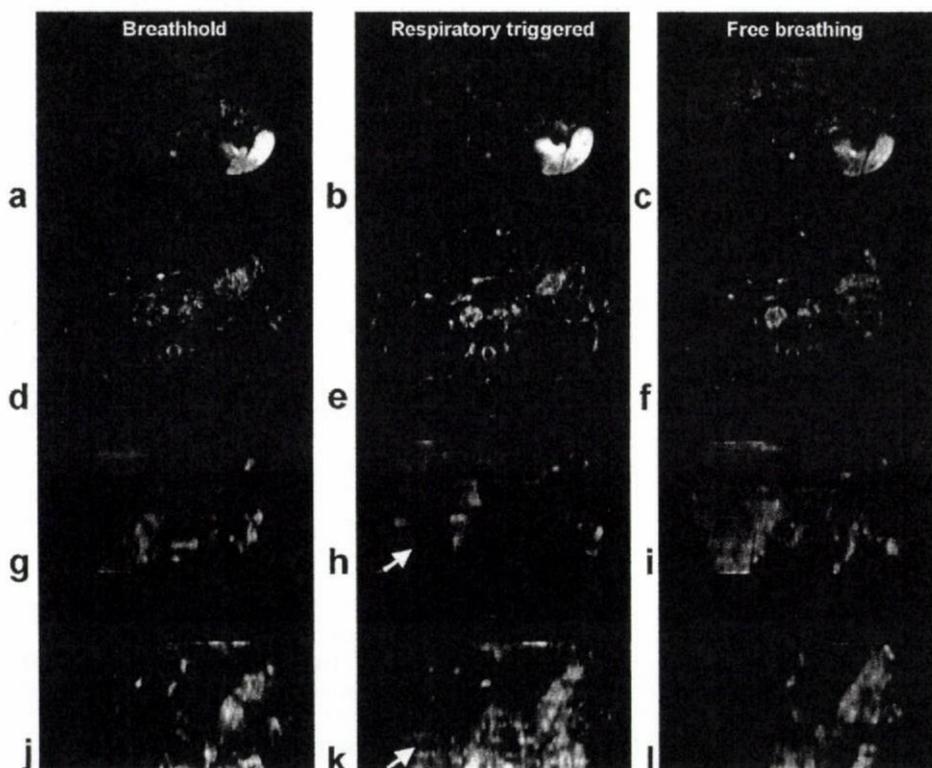

Figure 2. Representative example of diffusion-weighted images and ADC maps in breathhold, respiratory triggered, and free-breathing scanning (slice thickness/gap 7/1 mm), in one of the volunteers. A–C: Axial DWI scans at a b-value of 500 s/mm² and (D–F) corresponding ADC maps, at the level of the right portal vein. G–I: Coronally reformatted DWI scans at a b-value of 500 s/mm² and (J–L) corresponding ADC maps, at the level of the right portal vein. Note the relatively higher and more inhomogeneous ADCs in coronally reformatted respiratory triggered scans (H,K; arrows), compared to those in breathhold (G,J) and free-breathing (I,L) scans. This finding is less pronounced in the currently shown axial slices (A–F), because (inhomogeneities in) ADCs appear to vary in craniocaudal direction in respiratory triggered scanning in this volunteer (H,K).
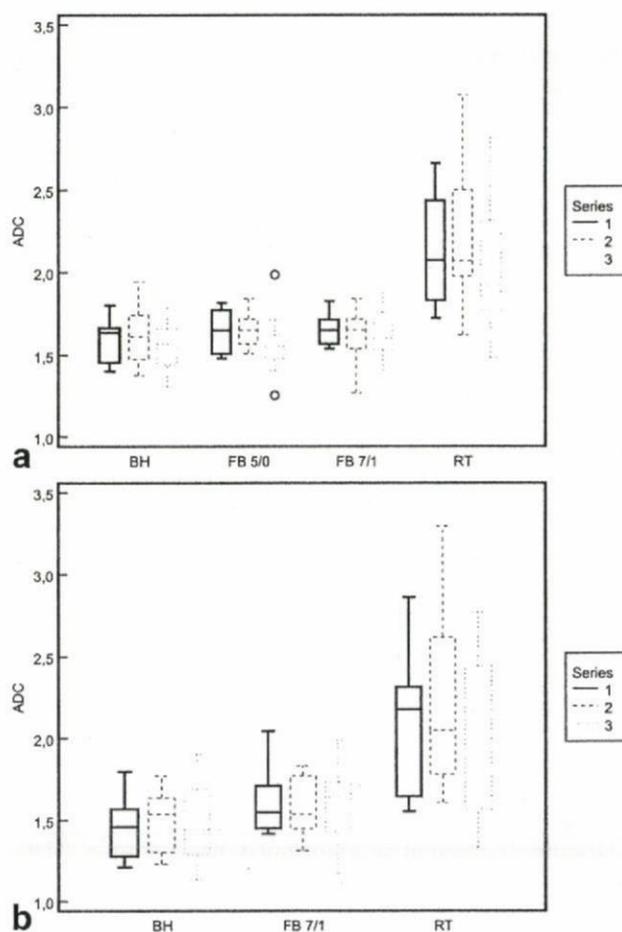
Figure 3. Box-and-whisker plots of ADC measurements of the normal liver parenchyma in **(A)** three series of coronally reformatted breathhold DWI (BH), free-breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), free-breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0), and respiratory triggered DWI (RT), and **(B)** three series of axial breathhold DWI (BH), free-breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and respiratory triggered DWI (RT).

gered, and free-breathing images (only with slice thickness/gap of 7/1 mm) at a b-value of 0 s/mm². One ROI was placed in the homogeneous liver parenchyma (visible vascular and biliary structures were excluded), at the lateral side of the right hepatic lobe (with at margin of at least 0.5 mm from the lateral border of the liver), at the slice through the right portal vein. The two other ROIs were placed four slice levels above and four slice levels below the slice through the right portal vein, respectively. Subsequently, all ROIs were copied and pasted on the I ADC maps, and ADCs were automatically calculated.

Statistical Analysis

First, liver isotropy in each of the coronally reformatted breathhold, respiratory triggered, and the two free-breathing DWI scans was evaluated by comparing ADCs obtained in three directions (P, M, and S). Next, I ADCs of the liver parenchyma among coronally refor-

matted breathhold, respiratory triggered, and the two free-breathing DWI scans were compared. Subsequently, I ADCs of the liver parenchyma among axial breathhold, respiratory triggered, and free-breathing DWI scans (only with slice thickness/gap of 7/1 mm) were compared, after averaging the measured ADCs of the three ROIs on each image. The nonparametric Kruskall–Wallis test was used for overall comparison and the nonparametric Wilcoxon's signed rank test was used for paired comparisons when the overall comparison was significant. Differences were considered significant when P values were less than 0.05.

Furthermore, reproducibility of ADC measurements on coronally reformatted breathhold, respiratory triggered, and the two free-breathing DWI scans was determined as mean absolute difference (bias) and 95% confidence interval of the mean difference (limits of agreement) according to the methods of Bland and Altman (8).

All statistical analyses were done three times, since three series of one breathhold, one respiratory triggered, and two free-breathing DWI scans were obtained in each volunteer. Statistical analyses were executed using Statistical Package for the Social Sciences v. 12.0 software (SPSS, Chicago, IL) and MedCalc Software (MedCalc, Mariakerke, Belgium).

RESULTS

Mean actual scan time of respiratory triggered DWI was 75.0 second (range, 41.7–117.1 sec). The liver parenchyma was isotropic in each of the coronally reformatted breathhold, respiratory triggered, and free-breathing DWI scans; only small, nonsignificant differences in ADCs were observed among P, M, and S directions (Table 1). On coronally reformatted images, ranges of mean ADCs in breathhold DWI, respiratory triggered DWI, free-breathing DWI with slice thickness/gap of 7/1 mm, and free-breathing DWI with slice thickness/gap of 5/0 mm were 1.57–1.62×10^{-3} mm²/sec, 2.07–2.27×10^{-3} mm²/sec, 1.62–1.65×10^{-3} mm²/sec, and 1.57–1.66×10^{-3} mm²/sec, respectively (Table 2). ADCs in respiratory triggered DWI were significantly higher ($P < 0.005$) than ADCs in breathhold and free-breathing DWI (both with slice thickness/gap of 7/1 mm and 5/0 mm) (Figs. 2, 3; Tables 2, 3). There were no significant differences in ADCs among breathhold and free-breathing DWI scans (Tables 2, 3). ADCs in respiratory triggered DWI were more scattered than ADCs in other sequences (Figs. 2, 3; Table 2). Analysis of axial images resulted in similar findings (Figs. 2, 3; Tables 4, 5).

Figure 4 shows the results of Bland–Altman reproducibility analysis of ADC measurements on coronally reformatted images of each sequence. In breathhold DWI, mean bias in ADC measurement between two scans varied from -0.02×10^{-3} mm²/sec to 0.05×10^{-3} mm²/sec, and the limits of agreement varied between $\pm 0.16 \times 10^{-3}$ mm²/sec and $\pm 0.24 \times 10^{-3}$ mm²/sec (Fig. 4; Table 6). In respiratory triggered DWI, mean bias in ADC measurement between two scans varied from -0.14×10^{-3} mm²/sec to 0.20×10^{-3} mm²/sec, and the limits of agreement varied between $\pm$

Table 3

P-values of Paired Comparisons of Isotropic ADCs of the Normal Liver Parenchyma

Series	BH vs. RT	BH vs. FB 7/1	BH vs. FB 5/0	RT vs. FB 7/1	RT vs. FB 5/0	FB 7/1 vs. FB 5/0
1	0.003	0.110	0.130	0.003	0.003	0.477
2	0.003	0.824	0.328	0.003	0.003	0.286
3	0.003	0.062	0.477	0.004	0.003	0.110

Coronally reformatted breathhold DWI (BH), respiratory triggered DWI (RT), free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and free breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0).

0.59×10^{-3} mm^2/sec and $\pm 0.60 \times 10^{-3}$ mm^2/sec (Fig. 4; Table 6). In free-breathing DWI with slice thickness/gap of 7/1 mm, mean bias in ADC measurement between two scans varied from -0.03×10^{-3} mm^2/sec to 0.03×10^{-3} mm^2/sec, and the limits of agreement varied between $\pm 0.20 \times 10^{-3}$ mm^2/sec and $\pm 0.29 \times 10^{-3}$ mm^2/sec (Fig. 4; Table 6). In free-breathing DWI with slice thickness/gap of 5/0 mm, mean bias in ADC measurement between two scans varied from -0.01×10^{-3} mm^2/sec to 0.09, and the limits of agreement varied between $\pm 0.21 \times 10^{-3}$ mm^2/sec and $\pm 0.29 \times 10^{-3}$ mm^2/sec (Fig. 4; Table 6).

DISCUSSION

ADC measurements may be helpful in characterizing liver lesions. In general, malignant lesions have lower ADCs, whereas benign lesions have higher ADCs, although variable overlap occurs between both groups (3–6). In addition, ADC measurements may be helpful in diagnosing and quantifying liver fibrosis (9,10). Furthermore, ADC measurements may be of value in predicting or assessing response to chemotherapy (11,12). This study is the first to compare and assess the reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing DWI.

First, our results indicate that measured diffusion in the normal liver parenchyma is isotropic, in both breathhold, respiratory triggered, and free-breathing DWI. These findings are concordant with those of Taouli et al (6), who showed that diffusion is isotropic in the liver parenchyma in breathhold DWI. However, Taouli et al (6) did not investigate liver isotropy in respiratory triggered and free-breathing DWI, as was done in the present study.

Second, the present study shows that ADCs of the normal liver parenchyma in breathhold DWI, free-breathing DWI with thick slice acquisition (slice thickness/gap: 7/1 mm), and free-breathing DWI with thin slice acquisition (slice thickness/gap: 5/0 mm) were all around 1.6×10^{-3} mm^2/sec, which again is in agreement with previous studies using breathhold DWI (6,13–15), and were not significantly different from each

other. In contrast, ADCs in respiratory triggered DWI were significantly higher, measuring $2.05–2.27 \times 10^{-3}$ mm^2/sec on average, and were more scattered. ADCs were initially measured on coronally reformatted images in order to prevent slice selection bias when measuring ADCs on axial images. A disadvantage of ADC measurements on coronally reformatted images is that some minimal partial volume averaging effects were introduced because of the use of a cubic spline interpolation algorithm to fill the 1-mm gaps between the 7-mm-thick slices. However, ADC measurements on axial slices are essentially more susceptible to partial volume averaging effects than those on coronal images because the largest displacement of the liver is in the craniocaudal direction during breathing. Nevertheless, ADC measurements were also done on axial images because the majority of previously published articles only analyzed axial images. ADC measurements on axial images were similar to those on coronally reformatted images; we may therefore state that our results on coronally reformatted images are reliable.

Third, our results show that reproducibility of ADC measurements of the liver parenchyma in breathhold and free-breathing DWI (both with thick and thin slice acquisition) was comparable and acceptable, considering the fact that mean bias in ADC measurement between two scans did not exceed $\pm 0.10 \times 10^{-3}$ mm^2/sec, and limits of agreement were lower than $\pm 0.30 \times 10^{-3}$ mm^2/sec. On the other hand, in respiratory triggered DWI, reproducibility was poor; maximum mean bias in ADC measurement between two scans was 0.20×10^{-3} mm^2/sec, with limits of agreement of up to $\pm 0.60 \times 10^{-3}$ mm^2/sec.

A possible explanation for the more scattered and less reproducible ADC measurements in respiratory triggered DWI is a mismatch in end-expiratory diaphragm levels between sequential triggering events. Consequently, slice levels of images obtained with different b-values may not be identical, which may influence ADC measurements. However, although the same effect can be expected in free-breathing DWI (slice levels of images obtained with different b-values may be different because of the allowance of respiratory motion),

Table 4

Overall Comparison of Isotropic ADCs (in 10-3 mm2/s) of the Normal Liver Parenchyma

Series	BH	RT	FB 7/1	P value
1	1.45 ± 0.21	2.07 ± 0.42	1.62 ± 0.19	0.001
2	1.49 ± 0.19	2.22 ± 0.53	1.58 ± 0.17	0.000
3	1.48 ± 0.24	2.06 ± 0.51	1.59 ± 0.25	0.016

Axial breathhold DWI (BH), respiratory triggered DWI (RT), and free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1).
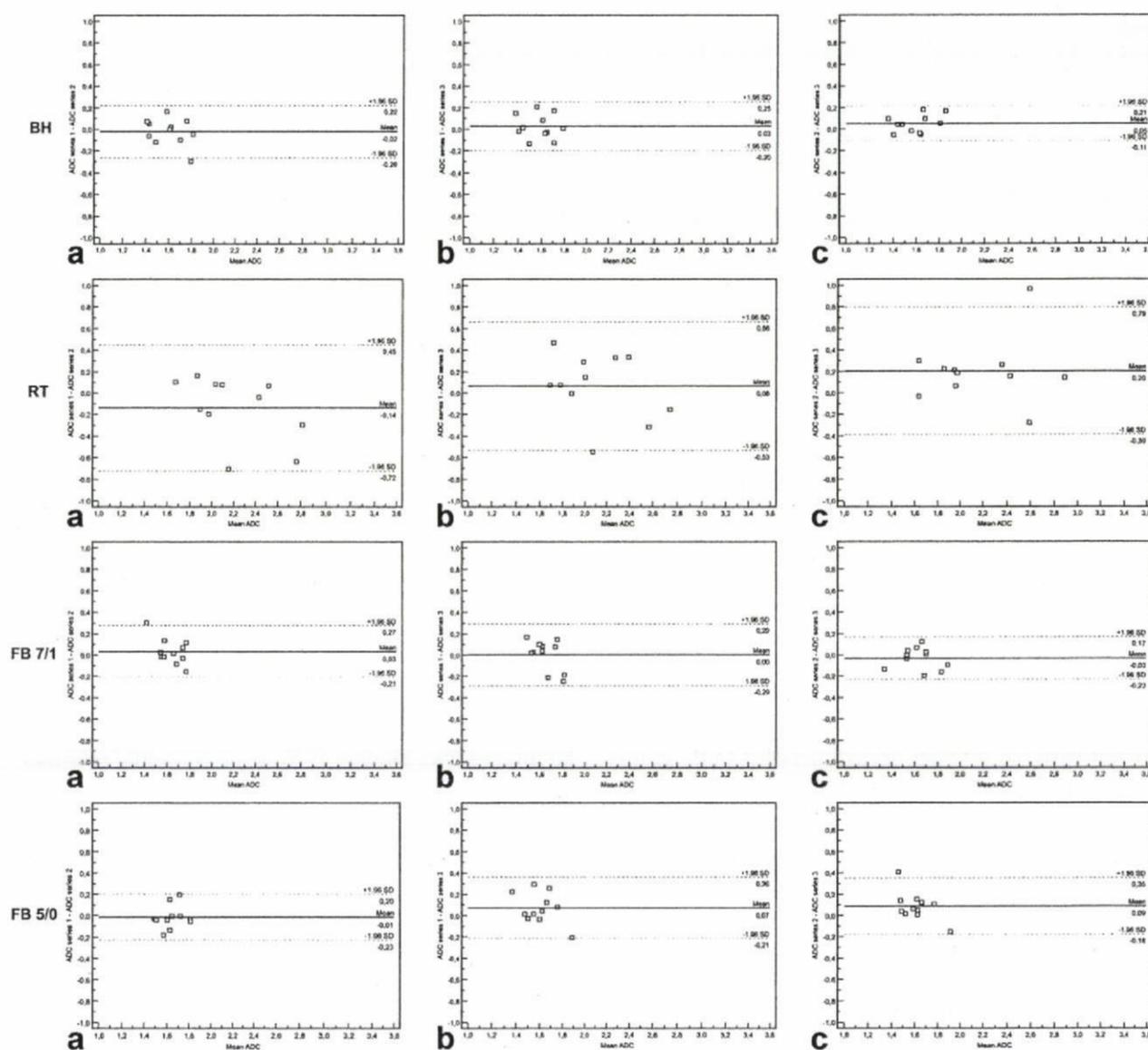
Figure 4. Reproducibility of ADC measurements of the normal liver parenchyma in coronally reformatted breathhold DWI (BH), respiratory triggered DWI (RT), free-breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and free-breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0). Bland–Altman plots of difference of ADC measurements (y-axis) against mean ADC measurement (x-axis), with mean absolute difference (bias) (continuous line) and 95% confidence interval of the mean difference (limits of agreement) (dashed lines) for **(a)** scans of series 1 and 2, **(b)** scans of series 1 and 3, and **(c)** scans of series 2 and 3.

ADCs in free-breathing DWI were comparable to those in breathhold DWI, and less scattered. Therefore, there has to be another explanation for the higher, more scattered, and less reproducible ADC measurements in respiratory triggered DWI, and this may be the so-called hepatic pseudo-anisotropy artifact (16). This artifact was recently described by Nasu et al (16), and consists of localized signal loss in the liver parenchyma, depending on the difference of MPG direction. Because the healthy liver is elastic and easily deformed by respiratory movement (17), it can be hypothesized that this artifact probably occurs because of intravoxel deformation and/or acceleration of the liver during the period in which the MPGs are applied, at the end of expiration.

Nasu et al (16) reported an overall incidence of hepatic pseudo-anisotropy of 48%. The incidence in pseudo-anisotropy did not significantly differ among different MPG directions (16). Mean ADC in areas with pseudo-anisotropy $(2.41 \pm 0.44 \times 10^{-3} \text{ mm}^2/\text{sec})$ was significantly higher $(P < 0.001)$ than that in areas without pseudo-anisotropy $(1.54 \pm 0.16 \times 10^{-3} \text{ mm}^2/\text{sec})$, and ADCs were more scattered in areas with pseudo-anisotropy (16). If pseudo-anisotropy would occur during the entire respiratory cycle, there would have been no difference in ADCs between respiratory triggered DWI (image acquisition at the end of expiration) and free-breathing DWI (image acquisition during the entire respiratory cycle). However, in our study, ADCs in free-

Table 5
P-values of Paired Comparisons of Isotropic ADCs of the Normal Liver Parenchyma

Series	BH vs. RT	BH vs. FB 7/1	RT vs. FB 7/1
1	0.004	0.110	0.006
2	0.003	0.050	0.004
3	0.010	0.286	0.004

Axial breathhold DWI (BH), respiratory triggered DWI (RT), and free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1).

breathing DWI were lower and less scattered than those in respiratory triggered DWI. This suggests that the incidence and/or magnitude of the pseudo-anisotropy artifact may be the highest at the end of the expiration. Nevertheless, the big difference in ADCs between respiratory triggered and free-breathing DWI cannot only be explained by the pseudo-anisotropy artifact. Another possibility that cannot be ruled out is that the higher ADCs obtained with (end-expiratory) respiratory triggered DWI in this study may result from a systematic higher blood flow rate at the end of expiration. It may therefore be of value to evaluate different phases of the respiratory cycle for respiratory triggered DWI, although this was not done in the present study. The applied respiratory triggering technique (a pressure sensor fixed to the hypochondrial region by an elastic belt was used in the present study) may also play a role, and the effect of other respiratory triggering techniques on ADC measurements, such as navigator-echo-based respiratory triggering, should be investigated further.

In a previous experience [18], actual scan time in respiratory triggered DWI was two to three times longer than that in free-breathing DWI, when using the same MR parameters. However, to take into account time-efficacy with regard to clinical implementation of DWI of the liver, we planned to obtain an equivalent actual scan time in respiratory triggered and free-breathing DWI in the present study. Despite our initial planning, mean actual scan time in respiratory triggered DWI (75.0 sec) was still more than scan time in free-breathing DWI with thick slice acquisition (51.7 sec) and free-breathing DWI with thin slice acquisition (64.7 sec). Therefore, ADC measurements of the liver parenchyma in respiratory triggered DWI are not only higher and less reproducible than in breathhold and free-breathing DWI, but it should also be realized that respiratory triggered DWI suffers either from a considerably longer actual scan time or from lower SNR when employing the same actual scan time.

A limitation of the present study is that only two b-values (0 s/mm² and 500 s/mm²) were used for calculation of ADCs, and overestimation of ADCs by including the perfusion fraction was possible [20]. However, the use of higher b-values would have diminished image quality. Another possible drawback of this study is that it was carried out on young, healthy volunteers. However, because these study participants were young and healthy, they were all very well able to hold their breath for breathhold DWI, and to perform normal, regular breathing for respiratory triggered and free-breathing DWI (as was checked by the operator on the monitoring screen). If we would have extended our

study to patients, who are likely to be older and less cooperative to carry out the breathing instructions, our results may have been less reliable. Another limitation of the present study is that comparison and assessment of reproducibility of ADC measurements were only done in the normal liver parenchyma. It is important to realize that the results may be different in hepatic and extrahepatic tumors, especially in free-breathing DWI. Free-breathing DWI theoretically causes significant blurring of the image. ADC measurements of small liver lesions in free-breathing DWI may therefore be affected by partial volume averaging effects with the surrounding liver parenchyma. Consequently, ADC measurements of liver lesions may be less accurate and less reproducible in free-breathing DWI. In a phantom study, Muro et al [19,20] reported that ADCs of slowly moving phantoms, simulating respiratory motion, were no more than 10% different from ADCs of static phantoms, if the ROIs were placed (and remained) within the homogeneous phantom. However, if the ROIs covered the edge of the phantom, ADCs were markedly different [19,20]. Thus, the results of their study indicate that ADC measurements of homogeneous tissue (eg, the liver parenchyma) appear to be accurate, whereas ADC measurements of heterogeneous tissue or small lesions (eg, small liver lesions) may be unreliable in free-breathing DWI. Nasu et al [21] compared ADCs of hepatic malignant tumors in both respiratory triggered DWI and free-breathing DWI. Thirty patients with liver metastasis or hepatocellular carcinoma in the right hepatic lobe were examined. The mean ADC in respiratory triggered DWI was 1.36×10^{-3} mm²/sec ± 0.33, whereas the mean ADC in free-breathing DWI was 1.47×10^{-3} mm²/sec ± 0.61. The statistical difference between both mean ADCs was not clear ($P = 0.050$). However, the individual ADCs in free-breathing DWI were more scattered than those in respiratory triggered DWI, which reflects the potential lower accuracy and reproducibility of ADC measurements of liver lesions in free-breathing DWI. Unfortunately, Nasu et al [21] did not investigate the reproducibility of ADC measure-

Table 6
Mean Absolute Difference (Bias) and 95% Confidence Interval of the Mean Difference (Limits of Agreement) of ADC Measurements (in 10^{-3} mm²/s)

Sequence	Series of scans	Mean bias	Limits of agreement
BH	1 and 2	−0.02	±0.24
	1 and 3	0.03	±0.23
	2 and 3	0.05	±0.16
RT	1 and 2	−0.14	±0.59
	1 and 3	0.06	±0.60
	2 and 3	0.20	±0.59
FB7/1	1 and 2	0.03	±0.24
	1 and 3	0.00	±0.29
	2 and 3	−0.03	±0.20
FB5/0	1 and 2	−0.01	±0.21
	1 and 3	0.07	±0.29
	2 and 3	0.09	±0.27

Two series of scans in coronally reformatted breathhold (BH), respiratory triggered (RT), free breathing DWI with slice thickness/gap of 7/1 mm (FB 7/1), and free breathing DWI with slice thickness/gap of 5/0 mm (FB 5/0).

ments of liver lesions. A precise measurement of ADCs is needed in both liver parenchyma and liver lesions. Therefore, more research is needed to determine the accuracy and reproducibility of ADC measurements in liver lesions, with use of breathhold, respiratory triggered, and free-breathing DWI.

In conclusion, the results of this study indicate that ADC measurements of the liver parenchyma in respiratory triggered DWI are significantly higher and less reproducible than in breathhold and free-breathing DWI. Therefore, it is our opinion that ADC measurements of the liver parenchyma should be performed by use of breathhold or free-breathing DWI.

REFERENCES

1. Low RN. Abdominal MRI advances in the detection of liver tumours and characterisation. Lancet Oncol 2007;8:525–535.
2. Naganawa S, Kawai H, Fukatsu H, et al. Diffusion-weighted imaging of the liver: technical challenges and prospects for the future. Magn Reson Med Sci 2005;4:175–186.
3. Gourtsoyianni S, Papanikolaou N, Yarmenitis S, Maris T, Karantanas A, Gourtsoyiannis N. Respiratory gated diffusion-weighted imaging of the liver: value of apparent diffusion coefficient measurements in the differentiation between most commonly encountered benign and malignant focal liver lesions. Eur Radiol 2008;18:486–492.
4. Bruegel M, Holzapfel K, Gaa J, et al. Characterization of focal liver lesions by ADC measurements using a respiratory triggered diffusion-weighted single-shot echo-planar MR imaging technique. Eur Radiol 2008;18:477–485.
5. Demir OI, Obuz F, Sağol O, Dicle O. Contribution of diffusion-weighted MRI to the differential diagnosis of hepatic masses. Diagn Interv Radiol 2007;13:81–86.
6. Taouli B, Vilgrain V, Dumont E, Daire JL, Fan B, Menu Y. Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients. Radiology 2003;226:71–78.
7. Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S, Van Cauteren M. Diffusion weighted whole body imaging with background body signal suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display. Radiat Med 2004; 22:275–282.
8. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.
9. Girometti R, Furlan A, Bazzocchi M, et al. Diffusion-weighted MRI in evaluating liver fibrosis: a feasibility study in cirrhotic patients. Radiol Med (Torino) 2007;112:394–408.
10. Taouli B, Tolia AJ, Losada M, et al. Diffusion-weighted MRI for quantification of liver fibrosis: preliminary experience. AJR Am J Roentgenol 2007;189:799–806.
11. Koh DM, Scurr E, Collins D, et al. Predicting response of colorectal hepatic metastasis: value of pretreatment apparent diffusion coefficients. AJR Am J Roentgenol 2007;188:1001–1008.
12. Liapi E, Geschwind JF, Vossen JA, et al. Functional MRI evaluation of tumor response in patients with neuroendocrine hepatic metastasis treated with transcatheter arterial chemoembolization. AJR Am J Roentgenol 2008;190:67–73.
13. Yoshikawa T, Kawamitsu H, Mitchell DG, et al. ADC measurement of abdominal organs and lesions using parallel imaging technique. AJR Am J Roentgenol 2006;187:1521–1530.
14. Oner AY, Celik H, Oktar SO, Tali T. Single breath-hold diffusion-weighted MRI of the liver with parallel imaging: initial experience. Clin Radiol 2006;61:959–965.
15. Taouli B, Martin AJ, Qayyum A, et al. Parallel imaging and diffusion tensor imaging for diffusion-weighted MRI of the liver: preliminary experience in healthy volunteers. AJR Am J Roentgenol 2004; 183:677–680.
16. Nasu K, Kuroki Y, Fujii H, Minami M. Hepatic pseudo-anisotropy: a specific artifact in hepatic diffusion-weighted images obtained with respiratory triggering. MAGMA 2007;20:205–211.
17. Rohlfing T, Maurer CR Jr, O'Dell WG, Zhong J. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. Med Phys 2004;31:427–432.
18. Takahara T, Ogino T, Okuaki T, et al. Respiratory gated body diffusion weighted imaging avoiding prolongation of scan time: tracking only navigator echo (TRON) technique. Presented at the Joint Annual Meeting ISMRM-ESMRMB, Berlin, Germany, May 19–25, 2007.
19. Muro I, Takahara T, Horie T, et al. Influence of respiratory motion in body diffusion weighted imaging under free breathing (examination of a moving phantom). Nippon Hoshasen Gijutsu Gakkai Zasshi 2005;61:1551–1558.
20. Koh DM, Takahara T, Imai Y, Collins DJ. Practical aspects of assessing tumors using clinical diffusion-weighted imaging in the body. Magn Reson Med Sci 2008;6:211–224.
21. Nasu K, Kuroki Y, Sekiguchi R, Nawano S. The effect of simultaneous use of respiratory triggering in diffusion-weighted imaging of the liver. Magn Reson Med Sci 2006;5:129–136.

MAJOR PAPER

High b-value Diffusion-weighted Imaging in Normal and Malignant Peripheral Zone Tissue of the Prostate: Effect of Signal-to-Noise Ratio

Kazuhiro Kitajima[1,2]*, Yasushi Kaji[1], Kagayaki Kuroda[3], and Kazuro Sugimura[2]

[1]Department of Radiology, Dokkyo Medical University
880 Kitakobayashi, Mibu, Shimotsuga-gun, Tochigi 321–0293, Japan
[2]Department of Radiology, Kobe University Graduate School of Medicine, Kobe, Japan
[3]Department of Human and Information Sciences, Faculty of Information and Electronics,
Tokai University, Hiratsuka, Japan
(Received February 6, 2008; Accepted March 26, 2008)

Purpose: To determine whether the apparent diffusion coefficient (ADC) obtained using a high b-value (2,000 s/mm²) is superior to that using a standard b-value (1,000 s/mm²) for discriminating malignant from normal peripheral tissue in the prostate.

Methods: Twenty-six patients with biopsy-proven prostate cancer underwent 1.5T magnetic resonance (MR) imaging including single-shot, echo-planar diffusion-weighted imaging (DWI) with repetition time/echo time, 3500/88 ms; 4-mm slice thickness; 1-mm interslice gap; 144×128 matrix; field of view, 250×250 mm; number of excitations, 10; and b-values, 0, 1,000, and 2,000 s/mm². For each patient, ADC values were obtained for malignant and normal tissue using b = 1,000 and 2,000 in a monoexponential model. Signal-to-noise (SNR) and contrast-to-noise (CNR) ratios in DWI were also evaluated.

Results: At b = 1,000, the mean ADC ($\times 10^{-3}$ mm²/s) for malignant tissue was 0.82 ± 0.27 (range 0.43–1.29) and for normal tissue, 1.69 ± 0.23 (1.31–2.18). At b = 2000, the mean ADC for malignant tissue was 0.61 ± 0.19 (0.30–0.94) and for normal tissue, 1.01 ± 0.14 (0.73–1.35). Significant ADC overlap between the malignant and normal tissue was recognized at b = 2000. As b-value increased, the mean SNR within malignant tissue decreased by 21.6%, and mean CNR decreased 17.3%.

Conclusions: Under the same imaging conditions, measuring ADC using a high b-value (2,000 s/mm²) in a monoexponential model has little diagnostic advantage over using the standard b-value (1,000 s/mm²) in discriminating malignant from normal prostate tissue.

Keywords: *apparent diffusion coefficient (ADC), diffusion-weighted imaging (DWI), high b-value, prostate cancer, signal-to-noise ratio (SNR)*

Introduction

Prostate cancer is the second most common cause of cancer-related death in men in most Western countries and poses a growing health problem. Given the incidence and mortality figures for prostate cancer, coupled with an aging population and incidence that rises steeply with age, there is a great deal of interest in accurately diagnosing and staging this particular disease. Accurate diagnosis is essential for correct treatment, impacts the patient, and has important cost implications.[1,2]

To increase diagnostic accuracy, numerous au-

thors have utilized magnetic resonance (MR) imaging to examine the prostate. In addition to conventional morphological imaging, new methodologies, such as spectroscopy, dynamic contrast-enhanced imaging, and diffusion-weighted imaging (DWI), are under investigation.[3,4]

DWI is sensitive to random thermal motion of water molecules in biologic tissues and has been used with apparent diffusion coefficient (ADC) mapping, which is derived from DWI, to assess malignant tissues in various organs. Water molecule diffusion is more restricted in malignant than normal tissue, probably because malignant tissue has greater cellular density and less extracellular space. Although some *in vivo* studies have demonstrated the usefulness of ADC for diagnosing pros-

*Corresponding author, Phone: +81-282-86-1111, Fax: +81-282-86-4940, E-mail: kazu10041976@yahoo.co.jp

tate cancer, the b-values used in those studies were equal to less than 1000 s/mm^2.[5-11] ADC tended to be lower in prostatic cancer tissue than in normal tissue in peripheral and transition zones. The recent advent of more powerful gradient hardware has permitted generation of greater b-values (up to 3,000 s/mm^2 on some commercial 1.5T MR units) that allow higher diffusion sensitivity, excluding perfusion and T$_2$-weighted effects. However, at very high b-value, noise figures increase and signal-to-noise (SNR) ratios decrease, so, we need to establish whether the decreased SNR diminishes the merits of MR imaging using very high b-value. In this project, we compared ADC and SNR of malignant and normal peripheral zone tissues using standard (1,000 s/mm^2) and high (2,000 s/mm^2) b-values to determine whether a high b-value is superior to the standard value for discriminating malignant from normal prostate tissue and to assess the dependence of measured ADC values on SNR.

Materials and Methods

Patients

Between October 2006 and May 2007, 45 consecutive patients with elevated levels of prostate-specific antigen (PSA) underwent pelvic MR imaging and subsequent systematic biopsy of the prostate. Among 32 of these patients with prostate cancer proven by transrectal ultrasound-guided biopsy, malignancy was also identified at the same position of the peripheral zone on MR imaging in 26 by 2 experienced radiologists (reader A with 7 years' and reader B with 19 years' experience reading prostate MR images). The median interval between MR examinations and biopsy was 19 days (12–29 days). The patients ranged in age from 58 to 84 years (mean, 71 years) and had serum PSA values ranging from 4.2 to 78.1 ng/mL (mean, 11.5 ng/mL). The hospital's review committee approved the study, and all participants gave written informed consent.

MR imaging technique

Before undergoing biopsy, all subjects were scanned on a 1.5T MR scanner (Magnetom Symphony Syngo Vision; Siemens AG, Erlangen, Germany) using a phased array pelvic coil for signal reception. The maximum gradient specifications were 30 mT/m for amplitude and 125 mT/m/ms for slew rate.

The entire prostate gland and seminal vesicle were covered by axial T$_1$-weighted spin-echo images (repetition time/echo time [TR/TE], 560/13 ms) in addition to axial and coronal T$_2$-weighted turbo-spin-echo images (TR/TE, 4500/100 ms; echo train length, 11). These conventional images were obtained with 4-mm slice thickness, 1-mm interslice gap, 250 × 250-mm field of vision (FOV), and 256 × 205 matrix. Gadolinium-enhanced dynamic imaging of volumetric interpolated breath-hold examination (VIBE) was performed in axial orientation (TR/TE, 5.5/2.6 ms; flip angle 90°; matrix 256 × 192; FOV 30 × 30 cm; slice thickness/gap 3/0 mm; number of excitations (NEX) 10; total acquisition time 20 s).In this dynamic scanning, precontrast baselines were sequentially obtained 40, 60, 80, and 180 s after bolus injection of 0.1 mmol/kg of gadopentetate dimeglumine (Magnevist; Nihon Schering, Osaka, Japan) at a rate of 3.0 mL/s. Axial DWI were obtained using a Stejskal-Tanner spin-echo echo-planar imaging (EPI) sequence (TR/TE, 3500/88 ms; flip angle 90°; NEX 10; b-value 0, 1000, and 2000 s/mm^2; readout band bandwidth 1184 Hz/pixel; matrix 144 × 128; FOV 250 × 250 mm; and slice thickness/gap 4/1 mm) covering the entire prostate and seminal vesicles. Following the acquisition with b = 0, motion-probing gradients (MPGs) were applied in 3 orthogonal orientations to calculate ADC, giving a scan time of 4 min 30 s. An acceleration factor of two was applied using the modified sensitivity encoding (mSENSE) parallel imaging technique.[12] In this DWI study, the voxel size was 1.7 × 2.0 × 4.0 mm, equaling 13.56 mm^3.

Image analysis
Standard reference

For each patient, among the lesions proven malignant by biopsy, 2 experienced radiologists defined as malignant those that showed focal rounded hypointensity relative to the surrounding parenchyma in the peripheral zone on T$_2$-weighted images and/or contrast enhancement in the early phase (40, 60, or 80 s after gadolinium injection) and washout in the delayed phase (180 s) on dynamic images. We excluded from analysis the lesions found malignant by biopsy that were undetectable on T$_2$-weighted images and/or gadolinium-enhanced dynamic images. Those proven benign by biopsy that showed hyperintensity in the peripheral zone on T$_2$-weighted images were defined as normal by the same radiologists and the three places were examined in each patient.
ADC value measurements

With reference to the T$_2$-weighted images and gadolinium-enhanced dynamic images, circular regions of interest (ROIs) were placed in the ADC maps to the boundaries of malignant (22–55 mm^2) and normal areas (> 40 mm^2) to avoid inclusion of

different tissue types in the evaluation of ADC values on machine's body of MR imaging equipment.

DWI was obtained by acquiring conventional T_2-weighted images while filtering out the signal from high-mobility water molecules. In this method, attenuation of the water signal from diffusion weighting is given by: $I/Io = \exp[-b\ ADC]$, where I and Io are the signal intensities in the presence and absence of the diffusion weighting gradients, ADC is the molecular diffusion coefficient, and b is the diffusion-weighted factor expressed as seconds per square millimeter. The ADC values were calculated for a pair of b-values: 0 and 1,000 s/mm², or 0 and 2,000 s/mm².

SNR and CNR measurements

The SNRs in the malignant and normal peripheral zones in each patient were calculated for DWIs with b-values of 1,000 and 2,000 s/mm². Because no background space was included in the relatively small FOV filled with the tissue signals, we used the mean of the standard deviation (SD) of the signals in the bilateral internal obturator muscles in place of those in the background.[13] Thus, the SNR and contrast-to-noise ratio (CNR) was defined as: $SNR = S/\delta_{noise}$, $CNR = (S_{normal} - S_{malignant})/\delta_{noise}$, where S is the average of the signal intensity in the malignant or normal peripheral zones and δ_{noise} is the mean SD for the signals in the bilateral internal obturator muscles. The circular ROIs placed within the muscles were consistently larger than 40 mm².

Data and statistical analysis

A total of 30 ROIs were placed in the malignant areas and 78 ROIs in the normal zones of the 26 patients. We compared ADCs using Mann-Whitney test and SNRs by paired-sample t-test. $P < 0.05$ was regarded as statistically significant.

Results

Of the 30 malignant lesions, 26 were located only in the peripheral zones, and four had invaded the transitional zones. The short axis diameter of the 30 malignant lesions ranged from 6 to 24 mm (mean 12.4 mm). The 30 lesions comprised 5 well-differentiated adenocarcinomas, 14 moderately differentiated adenocarcinomas, 4 poorly to moderately differentiated adenocarcinomas, and 7 poorly differentiated adenocarcinomas. Gleason's scores of the 30 lesions were 4 in 2 patients, 5 in 2, 6 in 3, 7 in 11, 8 in 5, and 9 in 7.

As shown in Fig. 1 and Table, the mean ADCs ($\times 10^{-3}$ mm²/s) for malignant tissues at b=1,000 s/mm² were 0.82 ± 0.27 (range 0.43–1.29) and for normal tissues 1.69 ± 0.23 (range 1.31–2.18) ($P = 1.11 \times 10^{-15}$, Mann-Whitney test). The corresponding values at b=2,000 s/mm² were 0.61 ± 0.19 (range 0.30–0.94) for malignant tissues and 1.01 ± 0.14 (range 0.73–1.35) for normal tissues ($P = 1.54 \times 10^{-12}$, Mann-Whitney test). At the higher b-value, the mean ADC decreased in malignant tissue by 25.6% and in normal peripheral zone tissue by 40.2%.

As shown in Fig. 2 and Table, the mean SNR for malignant tissue with b=1,000 was 48.7 ± 13.5 and with b=2,000 s/mm², 38.2 ± 11.9 ($P = 4.75 \times 10^{-11}$, paired t-test). The mean SNR for normal tissue with b=1,000 was 33.2 ± 7.9 and with 2,000 s/mm², 25.3 ± 6.7 ($P = 6.21 \times 10^{-8}$, paired t-test). At the higher b-value, the mean SNR decreased in malignant tissue by 21.6% and in normal peripheral zone tissue by 23.8%. The CNR with b=1,000 was 15.6 ± 8.1 and with 2,000 s/mm², 12.9 ± 7.5 ($P = 2.89 \times 10^{-6}$, paired t-test). As b-value increased, mean CNR decreased by 17.3%.

Figure 3 shows (a) T_2-weighted; (b) DWI (b=1,000 s/mm²); and (c) DWI (b=2,000 s/mm²) images of a typical subject with prostate cancer in the left peripheral zone, representing ADC and SNR in this malignant lesion ($ADC_{1000} = 0.80 \times 10^{-3}$ mm²/s; $SNR_{1000} = 46.7$; $ADC_{2000} = 0.59 \times 10^{-3}$ mm²/s; $SNR_{2000} = 36.7$).

Discussion

In our series of measurements, ADC values decreased in both malignant and normal peripheral zone tissues when a higher b-value was used. The cutoff levels of the mean ADC between malignant

Table. SNR, CNR, and ADC of malignant and normal peripheral tissue at b=1,000 and 2,000 s/mm²

	b=1,000 s/mm²		b=2,000 s/mm²	
	Cancer	Non-cancer	Cancer	Non-cancer
SNR	48.7 ± 13.5	33.2 ± 7.9	38.2 ± 11.9	25.3 ± 6.7
CNR	15.6 ± 8.1		12.9 ± 7.5	
ADC ($\times 10^{-3}$ mm²/s)	0.82 ± 0.27	1.69 ± 0.23	0.61 ± 0.19	1.01 ± 0.14

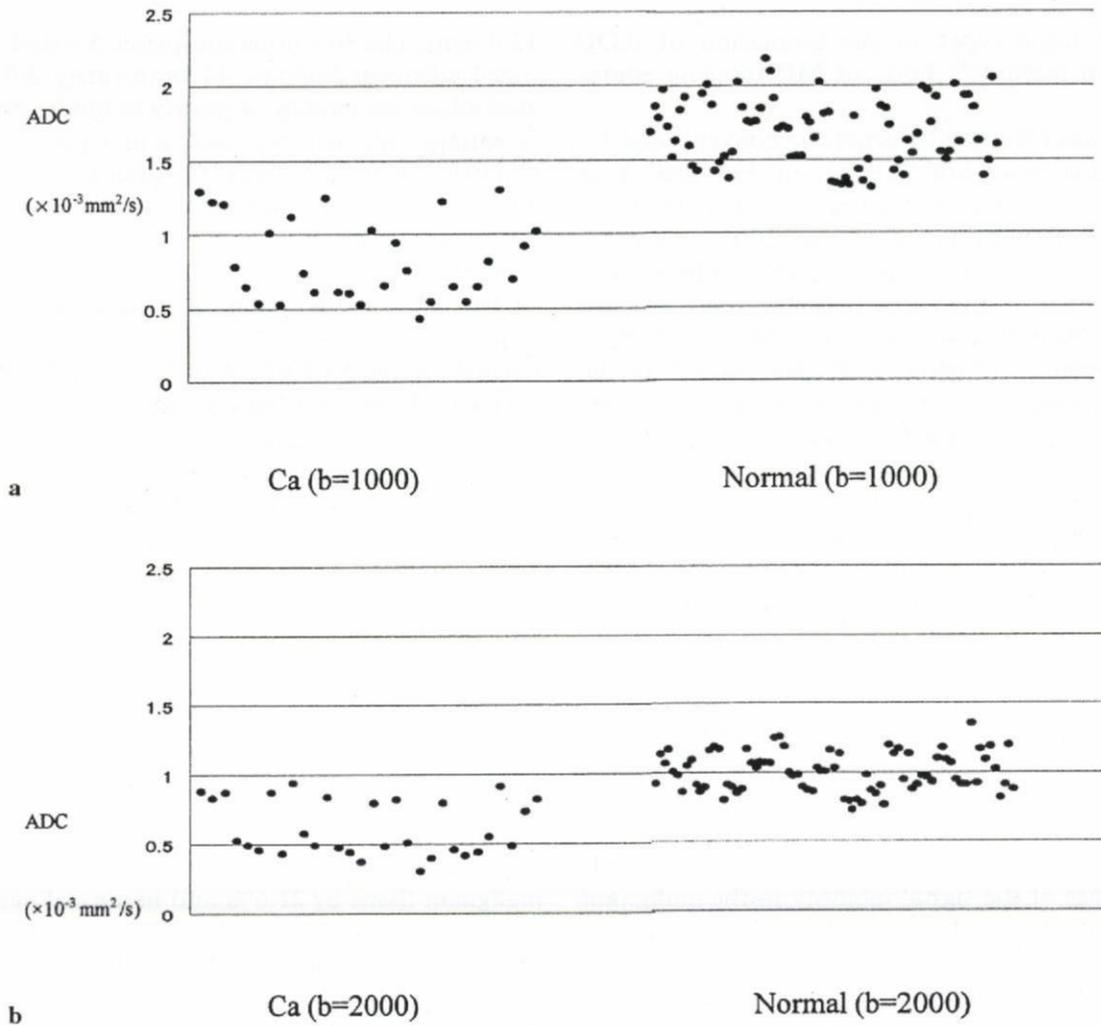**Fig. 1.** Plots show apparent diffusion coefficient (ADC) values of malignant and normal peripheral zone tissues ($\times 10^{-3}$ mm²/s) with (**a**) b = 1,000 and (**b**) 2,000 s/mm². ADC ($\times 10^{-3}$ mm²/s) for malignant tissue was 0.82 ± 0.27 (range 0.43–1.29) and for normal tissue, 1.69 ± 0.23 (range 1.31–2.18) ($P = 1.11 \times 10^{-15}$) at b = 1,000 s/mm². The corresponding values at b = 2,000 s/mm² were 0.61 ± 0.19 (range 0.30–0.94) for malignant tissue and 1.01 ± 0.14 (range 0.73–1.35) for normal tissue ($P = 1.54 \times 10^{-12}$).

and normal tissues ranged from 1.29 to 1.31×10^{-3} mm²/s at b = 1,000 s/mm². On the other hand, considerable overlap between the tissue ADC values was observed at b = 2,000 s/mm². In distinguishing malignant from normal tissue in the peripheral zone, ADC using a high b-value offered no apparent diagnostic advantage over that using the standard b-value. Although a number of studies using b-values equal to or smaller than 1,000 s/mm² have revealed considerable ADC overlap between malignant and normal tissue, Tanimoto's group[11] recognized no overlap based on total histological investigation after radical prostatectomy as a gold standard. They demonstrated that the mean ADC cutoff value between cancerous and noncancerous tissues ranged from 1.14 to 1.18×10^{-3} mm²/s.

Several reports have discussed whether DWI images and ADC maps using a high (2,000 or 3,000 s/mm²) compared to standard b-value (1,000 s/mm²) improve diagnostic accuracy for acute cerebral infarctions in human subjects.[14-18] Some have reported a high b-value superior to a standard b-value, and others have indicated no apparent diagnostic advantage using a high b-value. The latter group reports that increase in b-value results in the simultaneous decrease of the measured value of ADC and SNR in the DWI used for ADC calculation. This complicates interpretation of the ADC value; one cannot determine whether the ADC is truly small or has simply been misread as small because of the poor SNR. Delano and colleagues[14] demonstrated that an increase in b-value from

1,000 to 3,000 s/mm² made ADC decrease with an average of 30% for multiple regions. Burdette and associates[16] reported that the average SNR for b = 3,000 images decreased to 45% of that for b = 1,000 images and that the number of excitations should be increased by a factor of 4.7 to maintain an equivalent SNR. Although shortening TE and increasing NEX partially compensates for this decrease in SNR, increasing the NEX increases scanning time, and any motion during acquisition would degrade the images. In our series, the ADC map for b = 1,000 and 2,000 s/mm² was set at the

same TE and NEX at one time; in the future, it will be necessary to reexamine this using the same SNR and changing the TE and NEX. This decrease in SNR for higher b-value DWI may be partly overcome by applying a high-tesla MR system.

Some researchers demonstrated that when SNR was insufficient, ADC was highly SNR-dependent as a result of a combination of noise and choice of noise thresholding and that when SNR was sufficiently high, ADC was largely independent of SNR.[19-21] O'Halloran's group[19] demonstrated the effects of SNR on ADC measurements using porous phantoms and the lungs of human subjects using diffusion-weighted hyperpolarized helium–3 MR imaging and suggested that one approach for reducing the SNR-dependent bias in ADC measurement was to apply a fixed noise threshold to both the diffusion-weighted and diffusion-unweighted images.

In our series, we used a monoexponential model for water diffusion decay, but this model is inadequate when a large and high range of b-values is sampled, and biexponential analyses would be better suited for characterizing human brain[22-27] and prostate[28] diffusion decay curves. It is accepted that water in biological tissue can be described in terms of a fast equilibration between 2 main components–fast-diffusing water molecules, which are mostly extracellular, and slow-diffusing water molecules, which are either bound to macromolecules or confined within the cell membrane.[12,22-28] The diffusion of water molecules in the intracellular component has been shown to be an order of magnitude smaller than in the extracellular space, and the MR signal from the 2 classes of water molecules can be differentiated by using high-b-value diffu-
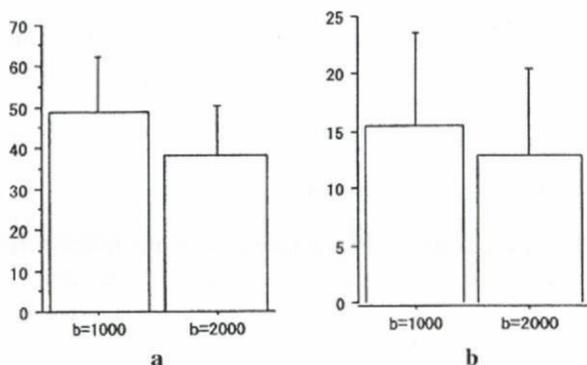

Fig. 2. Bar graph of (a) signal-to-noise ratios (SNR) within malignant peripheral zone tissues and (b) contrast-to-noise ratios (CNR) (mean ± standard deviation [SD]) for the diffusion-weighted images performed at b = 1,000 and 2,000 s/mm². The mean SNR for malignant tissue with b = 1,000 s/mm² was 48.7±13.5 and with 2,000 s/mm², 38.2±11.9 ($P = 4.75 \times 10^{-11}$). The CNR with b = 1,000 s/mm² was 15.6±8.1 and with 2,000 s/mm², 12.9±7.5 ($P = 2.89 \times 10^{-6}$).
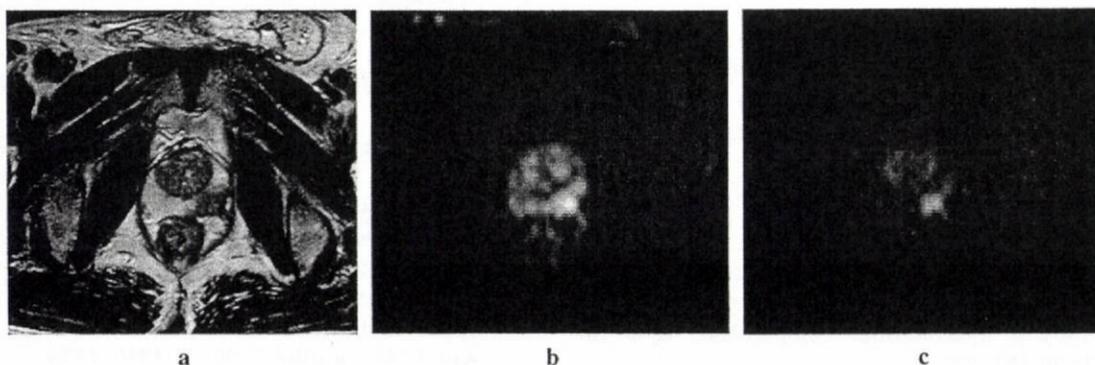

Fig. 3. (a) T$_2$-weighted image, (b) diffusion-weighted image (DWI) (b = 1000 s/mm²), (c) DWI (b = 2000 s/mm²) of a typical subject with prostate cancer in the left peripheral zone (arrow), representing apparent diffusion coefficient (ADC) and signal-to-noise ratio (SNR) in this malignant lesion (ADC$_{1000}$ = 0.80×10^{-3} mm²/s; SNR$_{1000}$ = 46.7; ADC$_{2000}$ = 0.59×10^{-3} mm²/s; SNR$_{2000}$ = 36.7).