Original Research

Inter- and Intra-rater Reproducibility of Quantitative Dynamic Contrast Enhanced MRI Using TWIST Perfusion Data in a Uterine Fibroid Model

Matthew S. Davenport, MD,[1*] Tobias Heye, MD,[2] Brian M. Dale, PhD,[2] Jeffrey J. Horvath, MD,[2] Steven R. Breault, MD,[2] Sebastian Feuerlein, MD,[3] Mustafa R. Bashir, MD,[2] Daniel T. Boll, MD,[2] and Elmar M. Merkle, MD[4]

Purpose: To determine the reproducibility of TWIST-derived (Time-Resolved Angiography with Interleaved Stochastic Trajectories) quantitative dynamic contrast enhanced (DCE) MRI in a uterine fibroid model.

Materials and Methods: The institutional review board approved this retrospective study. Dynamic contrast-enhanced TWIST datasets from 15 randomly selected 1.5 Tesla pelvic MR studies were postprocessed. Five readers recorded kinetic parameters (K^{trans} [volume transfer constant], v_e [extracellular extravascular space volume], k_{ep} [flux rate constant], iAUC [initial area under the gadolinium-time curve]) of the largest uterine fibroid using three region-of-interest (ROI) selection methods. Measurements were randomized and repeated three times, and measures of reproducibility were calculated.

Results: The intra-rater coefficients of variation (CVs, brackets indicate 95% confidence intervals) varied from 4.6% to 7.6% (K^{trans} 7.6% [6.1%, 9.1%], k_{ep} 7.2% [5.9%, 8.5%], v_e 4.6% [3.8%, 5.4%], and iAUC 7.2% [6.1%, 8.3%]). v_e was the most reproducible ($P < 0.05$). Inter-rater reproducibility was significantly ($P < 0.05$) greater for the large ROI method (range of intraclass correlation coefficients [ICCs] = 0.80–0.98 versus 0.48–0.63 [user-defined ROI] versus 0.41–0.69 [targeted ROI]). The uterine fibroid accounted for the greatest fraction of variance for the large ROI method (range across kinetic parameters: 83–98% versus 56–69% [user-defined ROI] versus 47–74% [targeted ROI]). The reader accounted for the greatest fraction of variance for the user-defined ROI method (0.4–14.1% versus 0.1–3.0% [large ROI] versus <0.1–1.5% [targeted ROI]).

Conclusion: Changes in TWIST-derived DCE-MRI kinetic parameters of up to 9–15% may be attributable to measurement error. Large DCE-MRI regions of interest are the most reproducible across multiple readers.

Key Words: DCE-MRI; perfusion; reproducibility
J. Magn. Reson. Imaging 2013;38:329–335.

NONINVASIVE BIOMARKERS THAT can be used to monitor drug efficacy or disease status have gained increasing attention in the scientific and medical communities. MRI is an attractive imaging modality because it provides a combination of qualitative and quantitative information. The ability to predict biologic behavior with minimal side effects is enticing, but this first requires validation to show that it is reproducible and that the biologic effect size is greater than the error inherent in the measurement.

Quantitative dynamic contrast-enhanced MRI (DCE-MRI) is a method of indirectly measuring the perfusion characteristics of a given tissue by acquiring repeated T1-weighted images following the intravenous administration of a gadolinium-based contrast material. Several commercial and public postprocessing tools are available that allow one to characterize the results of this data acquisition in the form of multiple kinetic parameters using a two-compartment Tofts paradigm (1): K^{trans} (volume transfer constant between plasma and extracellular extravascular space [EES]), v_e (EES volume), and k_{ep} (flux rate constant between EES and plasma). The iAUC (initial area under the gadolinium curve at 60–90 s) has also been considered a meaningful indicator of tissue perfusion (2–4). Many authors have suggested that use of these parameters may allow longitudinal monitoring of tumor biology (2–9).

However, DCE-MRI has considerable potential for systematic error. There are numerous factors that can affect the quantitative output. Therefore, data obtained at different sites or times may be difficult if not impossible to compare. We sought to assess the reproducibility of DCE-MRI kinetic parameter

[1]University of Michigan Health System, Ann Arbor, Michigan, USA.
[2]Duke University Medical Center, Durham, North Carolina, USA.
[3]University of Virginia Health System, Charlottesville, Virginia, USA.
[4]Basel University, Basel, Switzerland.
Contract grant sponsor: Xxxxxxx; Contract grant number: xxxxxxxx.
*Address reprint requests to: M.S.D., 1500 E. Medical Center Drive, Department of Radiology B2 A209P, University of Michigan Health System, Ann Arbor MI 48109. E-mail: matdaven@med.umich.edu
Received January 24, 2012; Accepted November 2, 2012.
DOI 10.1002/jmri.23974
View this article online at wileyonlinelibrary.com.

measurements when they are performed under controlled conditions, with as many sources of variation removed as possible (e.g., inputting the same DCE dataset into the same postprocessing workstation and calculating kinetic parameters using the same postprocessing algorithm and an identical arterial input function). We wanted to test the null hypothesis that under tightly controlled conditions there would be minimal variation within and between DCE-MRI readers. The purpose of our study was to determine the reproducibility of TWIST-derived quantitative DCE-MRI in a uterine fibroid model.

MATERIALS AND METHODS

Before the initiation of this investigation, Institutional Review Board approval was obtained. The study was carried out in compliance with the Health Insurance Portability and Accountability Act (HIPAA, USA). Patient informed consent was not required based on institutional policy and the retrospective nature of this investigation.

Subjects

All (N = 34) 1.5 Tesla (T) DCE pelvic MR studies performed before possible uterine fibroid embolization from August 21, 2009, to December 17, 2010, were identified by a retrospective review of the institutional electronic medical record system. None of the patients had previously undergone uterine artery embolization. A power calculation was performed prospectively to determine sample size adequacy. Effect size and standard deviation were estimated using nonpublished institutional pilot data (a significance level of 0.05 and statistical power of 0.8 were used for all calculations). The following assumptions were used: K^{trans} (standard deviation [SD] 0.3, effect size 0.3, needed sample size = 10), v_e (SD 0.2, effect size 0.3, needed sample size = 6), iAUC (SD 15, effect size 15, needed sample size = 10). Based on these calculations and assumptions, a conservative sample size of 15 was selected.

Of the 34 1.5T DCE pelvic MR studies retrieved in our initial query, four studies were excluded because they were performed on a unique MR system, one was excluded for erroneous capture (it was actually a pelvic MR venogram), one was excluded because it lacked a variable flip angle T1 map, and one was excluded because the largest uterine fibroid was too small to adequately measure (<1 cm in size). This left a total of 27 appropriate DCE pelvic MR studies with variable flip angle T1 maps performed on identical 1.5T systems (Siemens Avanto, Erlangen, Germany) and enhanced with 0.1 mmol/kg gadobenate dimeglumine (Bracco Diagnostics, Princeton, NJ) injected intravenously at 2 mL/s followed by a 20-mL saline flush at 2 mL/s. The mean volume of administered intravenous contrast material was 14.9 mL (range, 10–20 mL). Fifteen of these 27 studies were randomly selected based on the power calculation using a pseudorandom number generator in Microsoft Excel (Microsoft, Redmond, WA). These 15 studies comprised the study group.

Imaging Protocol

All DCE pelvic MR imaging studies were performed on the same 1.5T system (Siemens Avanto, Erlangen, Germany) in supine position with two anterior 6-element body phased-array coils and two 3-element posterior spine array coils. Background tissue T1 relaxation times were measured using a previously described variable flip angle technique with a three-dimensional (3D) spoiled gradient recalled echo sequence at 2, 8, and 20 degrees, respectively (10). Scanning parameters were as follows: echo time (TE) 1.31 milliseconds (ms); repetition time (TR) 3.74 ms; matrix 256 × 205; FOV 400 × 400 millimeter (mm); slice thickness 3–4 mm; phase encoding steps 153; echo train length 1; number of excitations (NEX) 4; parallel imaging acceleration factor 2. Dynamic contrast enhancement was imaged using a 4D, time resolved MR angiography sequence with interleaved stochastic trajectories (TWIST). A total of 30 consecutive measurements with a mean temporal resolution of 6.2 s (range, 5.3–9.0 s) resulted in a mean temporal coverage of 3.1 min (range, 2.7–4.5 min). Except for NEX 1, flip angle 12 and phase encoding steps 163; all other sequence parameters were matched exactly to the T1 map. T1 mapping and DCE imaging were performed in the coronal orientation to minimize motion artifact. The number of slices acquired per sequence was individually adjusted between subjects to the volume of interest and ranged from 48–80. Within each subject, the number of slices per temporal series was constant.

Perfusion Data Postprocessing

All DCE MRI studies were postprocessed (Tissue 4D, Siemens Healthcare) to obtain the quantitative perfusion parameters described above. The software calculated a voxel-wise Levenberg Marquardt fit to a two-compartment Tofts model (11), as well as the iAUC. Those voxel-wise parameters were then averaged over the selected ROI to obtain the quantitative parameters for the ROI.

A region of interest was placed in either the left common iliac artery or left external iliac artery on a representative coronal image to acquire a manual AIF. The artery with the least motion artifact was selected. An ROI was recorded for each dataset before quantitative analysis. The signal intensity over time curve was plotted to determine the contrast arrival time for each case. The contrast arrival time was determined by the postprocessing software, which analyzed the shape of the time signal intensity curve generated by the region of interest drawn in the inflow vessel (12). This generated a mean contrast arrival time of 0.33 min (range, 0.24–0.43 min). The contrast arrival time was referenced by the postprocessing workstation to access a library of population-averaged bi-exponential AIFs. An intermediate model AIF was selected for analysis in all cases from among the options of: fast, intermediate

and slow. The contrast arrival times for each of the datasets were recorded on the readers' data collection worksheet. This allowed each of the five readers to use the same AIF model.

Regions of Interest

Three two-dimensional region-of-interest (ROI) selection methods were chosen to test the variability in different types of measurements that might be performed in clinical practice. These regions of interest were directed toward the largest uterine fibroid in each case, which was preidentified for the readers by slice position and a written anatomic description (see Data Worksheet and Data Collection, below). These ROIs included: (i) a user-defined ROI with instructions to assess "the most enhancing component" of the target uterine fibroid (using a combination of early time point and kinetic parametric mapping image review), (ii) a large ROI encompassing at least three-quarters of the target uterine fibroid on the largest coronal section, and (iii) a targeted ROI of user-determined size directed to the target uterine fibroid at a specific location and slice position (Fig. 1). A specific size for the user-defined and targeted ROIs was not prescribed. The readers were instructed that the ROIs should not extend outside the margins of the target uterine fibroid and they were cautioned to avoid partial volume effects.

Data Worksheet and Data Collection

A data collection worksheet was created for each reader that targeted in advance the largest uterine fibroid in each patient. The location of each fibroid was specified using a representative slice location and series number, as well as a written anatomic description. For each patient, the following data were also specified: (i) the dose of intravenous contrast material administered (0.1 mmol/kg; range: 10–20 mL), (ii) the type of contrast material (gadobenate dimeglumine, Bracco Diagnostics, Princeton, NJ), and (iii) the AIF contrast arrival time.

Each combination of ROI method, uterine fibroid, and kinetic value (K^{trans}, k_{ep}, v_e, and iAUC) was measured and recorded three separate times by each of the five (MSD, TH, JJH, SRB, SF) readers (i.e., 36 values per fibroid per reader). The order of ROI placement was randomized on the data collection worksheet using a pseudorandom number generator in Microsoft Excel (Microsoft, Redmond, WA). Each set of three unique ROI selection methods (large ROI, user-defined ROI, and targeted ROI) was completed before the subsequent set was begun. After each set of unique ROI selection methods was completed, all ROIs were deleted from the images before the next set of was begun.

Before initiating data analysis, all readers were familiarized with the postprocessing workstation, trained to identify the target lesions by the methods described above, and practiced using postperfusion processing on nonstudy cases using the same instructions as they used for the study group. All readers were evaluated before study initiation by an experienced reader in MR perfusion analysis (blinded for review) to ensure competency with the task. All readers were familiar with the appearance of uterine fibroids and knew how to identify the target uterine fibroid using the localization information on the worksheet.

Data Analysis

Intra-rater reproducibility was assessed by calculating the coefficient of variation (CV = standard deviation [SD] / mean) for each set of three repeated ROI measurements. The estimated measurement error was determined by multiplying the mean CVs by 1.96. To determine if the CVs were influenced by any of the tested variables (reader, uterine fibroid, ROI method), a stepwise analysis of variance (ANOVA) was performed for all main effects using the minimum Bayes information criterion to select the best model.

Inter-rater agreement was directly assessed by calculating intraclass correlation coefficients (ICC) between readers for each ROI method and kinetic parameter. Inter-rater agreement was also assessed using a random-effect ANOVA to determine if any of the tested effects had an effect on the mean perfusion values measured and to determine the proportion of variance explained by each effect. In all cases, the modeled effects included: ROI method, reader, and individual uterine fibroid. The readers and uterine fibroids were considered random effects, and the ROI method was considered a fixed effect.

One K^{trans} data point and three k_{ep} data points out of 2700 total data points were excluded from analysis because they were extreme outliers (8–11 SD away from the mean) and believed to most likely represent transposition errors.

Fibroid volume was calculated using the following formula:

$$[\text{length} \times \text{width} \times \text{height} \times 0.52].$$

All statistical analysis was performed (BD) using JMP v 9.0 (SAS, Cary, NC). Numbers in brackets indicate 95% confidence intervals.

RESULTS

There were 15 uterine fibroids in 15 female patients with a mean age of 44 years (range, 28–60 years), a mean height of 1.65 meters (range, 1.52–1.83 m), and a mean weight of 75.8 kg (range, 54.4–136.5 kg). The mean uterine fibroid volume was 211 cm^3 (range, 1.49–637 cm^3). The mean maximum uterine fibroid diameter was 7.6 cm (range, 1.7–11.9 cm). Two uterine fibroids were located within the lower uterine segment, four were in the uterine body, and nine were in the uterine fundus. Four uterine fibroids were characterized as subserosal and 11 were characterized as intramural. One fibroid was extensively necrotic, with greater than 75% nonenhancing volume; five fibroids were moderately necrotic, with approximately 25–75% nonenhancing volume; one fibroid was mildly

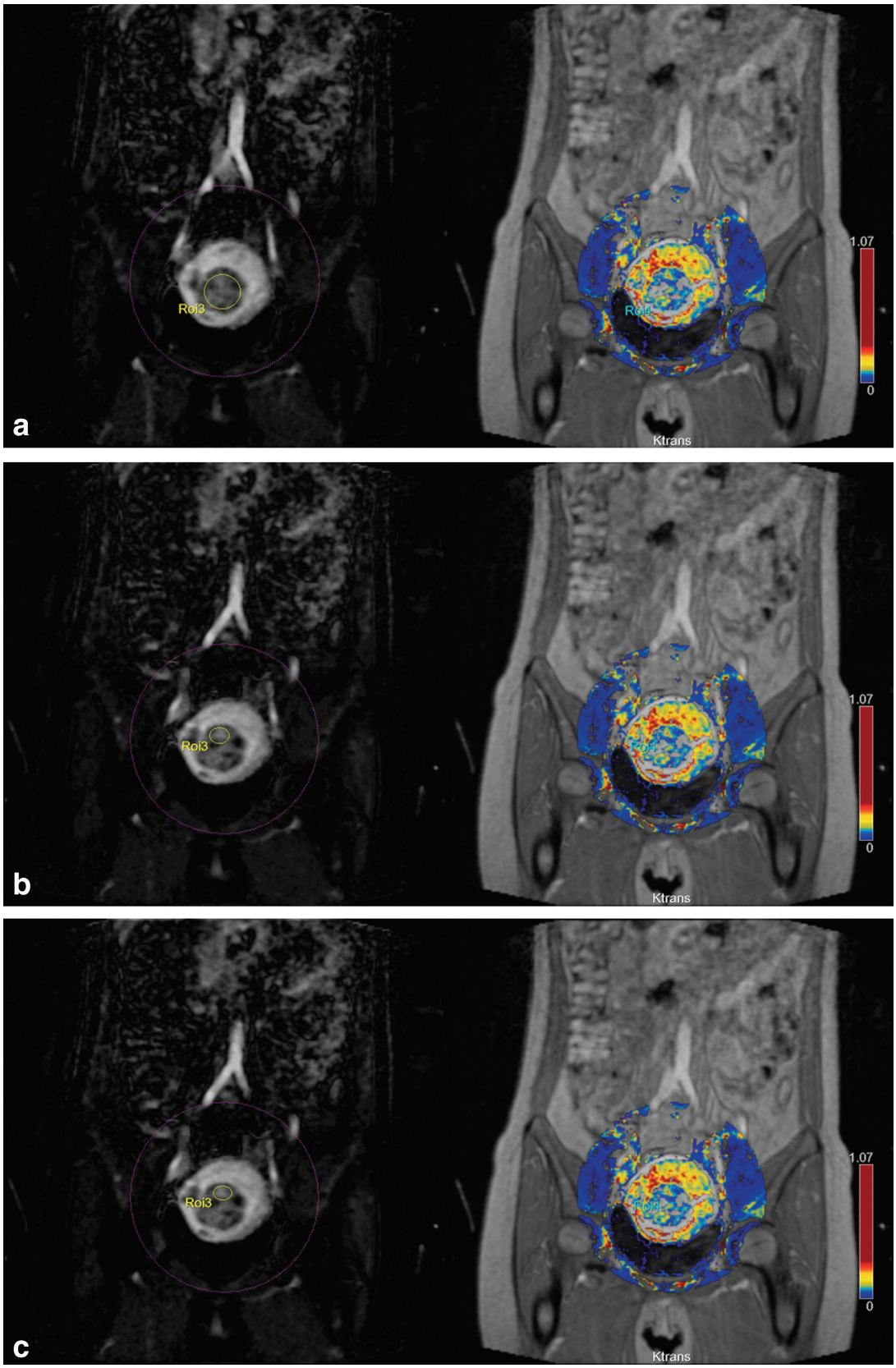

Figure 1. a–c: Representative T1-weighted dynamic contrast-enhanced subtraction images and nonsubtracted parametric maps (K^{trans} is mapped in these examples) for each of the three region-of-interest (ROI) types used during data collection: (a) large ROI, (b) user-directed ROI, and (c) targeted ROI.

Table 1
Intra-rater Reproducibility Expressed as the Coefficient of Variation (CV) for Each Kinetic Parameter and Region of Interest (ROI) Selection Method*

	Intra-rater CVs		
	Large ROI	Targeted ROI	User-defined ROI
iAUC	7.0% [4.6%, 9.4%]	6.9% [5.3%, 8.4%]	7.6% [5.9%, 9.3%]
K^{trans}	5.7% [4.3%, 7.0%]	8.9% [5.2%, 12.5%]	8.3% [6.0%, 10.5%]
K_{ep}	5.0% [3.5%, 6.5%]	6.5% [4.8%, 8.3%]	10.2% [7.1%, 13.3%]
v_e	3.9% [2.8%, 5.1%]	4.2% [3.1%, 5.3%]	5.6% [3.6%, 7.5%]

*The brackets indicate 95% confidence intervals.

necrotic, with less than 25% nonenhancing volume; and eight fibroids showed no measurable necrosis.

The mean CVs and 95% confidence intervals for intra-rater kinetic parameter reproducibility were: K^{trans} 7.6% [6.1%, 9.1%], k_{ep} 7.2% [5.9%, 8.5%], v_e 4.6% [3.8%, 5.4%], and iAUC 7.2% [6.1%, 8.3%]. Based on these mean CVs, the quantitative DCE measurement error is estimated at 9–15%. The iAUC CV was significantly ($P = 0.0002$) influenced by the reader, but accounted for a relatively small fraction of the total variance with an adjusted R^2 of 0.08. The K^{trans} and k_{ep} CVs were not strongly influenced by any of the tested factors according to the Bayes minimum information criterion. The individual intra-rater CVs for each combination of ROI selection method and kinetic parameter are displayed in Table 1.

Inter-rater reproducibility of all kinetic parameters as expressed by ICCs was significantly ($P < 0.0001$) greater for the large ROI method (ICCs = 0.80–0.98) compared with the user-defined (ICCs = 0.48–0.63) and targeted (ICCs = 0.41–0.69) ROI methods. Table 2 lists the ICCs for each kinetic parameter and ROI selection method. The large ROI method had the highest agreement between readers, indicating that encompassing a large fraction of the mass/uterine fibroid in the ROI has the highest inter-rater reproducibility. Although the user-defined ROI had the greatest intra-rater agreement, it had the lowest inter-rater agreement.

Table 3 lists the fraction of variance explained for each kinetic parameter by uterine fibroid, reader, and random error. The fraction of variance explained by differences in the measured uterine fibroid (as opposed to the reader or unexplained error) was greatest using the large ROI method (83–98%) compared with the user-defined (56–69%) and targeted (47–74%) ROI methods. The reader contributed the greatest fraction of variance when the user-defined ROI was used (0.4–14.1% versus 0.1–3.0% [large ROI] versus <0.1–1.5% [targeted ROI]).

The large ROI method had significantly lower ($P < 0.0001$) iAUC values (17.2 mmol/s [16.5, 17.8]) compared with the targeted (22.6 mmol/s [22.0, 23.3]) and user-defined ROIs (24.7 mmol/s [24.0, 25.3]), K^{trans} values ($P < 0.0001$, 0.32/min [0.31, 0.34], targeted = 0.45/min [0.44, 0.47], user-defined = 0.52/min [0.50, 0.53]), and v_e values ($P < 0.0001$, 0.54 mL [0.53,0.55], targeted = 0.59 mL [0.58,0.61], user-defined = 0.62 mL [0.61,0.63]). The k_{ep} values were not significantly different ($P = 0.86$) for the different ROI methods (0.95 [0.91,1.00]).

DISCUSSION

Noninvasive imaging-based in vivo monitoring of tumor biology and drug efficacy offers the potential to change the medical paradigm by individualizing oncologic therapy. However, before such an approach can be used in a widespread manner, it is important that a thorough analysis be first performed on the imaging modality itself to quantify the inherent systematic error. If an imaging biomarker suffers from a lack of reproducibility—within readers, between readers, or between sites—then results gleaned from these imaging studies will be difficult to interpret and, therefore, difficult to place in a clinical context.

DCE-MRI is one such biomarker that holds great promise but suffers from several confounding factors. Results obtained at one time point may be rendered incomparable to those obtained at a different time point if even one among a whole variety of factors have been changed in the interval (e.g., TR, flip angle, radiofrequency transmit field (B1), 2D versus 3D sequence type, temporal resolution, contrast material dose and injection rate, saline flush dose and injection rate, patient body size, patient body position, patient motion, proximity of the organ of interest to the magnet isocenter, matrix size, field strength, magnet manufacturer, postprocessing algorithm, arterial

Table 2
Inter-rater reproducibility expressed as the intraclass correlation coefficient (ICC) between raters for each combination of ROI selection method and kinetic parameter*

	Inter-rater ICCs		
	Large ROI	Targeted ROI	User-defined ROI
iAUC	0.80 [0.78, 0.82]	0.68 [0.64, 0.71]	0.63 [0.59, 0.67]
K^{trans}	0.88 [0.86, 0.89]	0.69 [0.65, 0.72]	0.62 [0.58, 0.66]
K_{ep}	0.98 [0.98, 0.99]	0.41 [0.36, 0.47]	0.48 [0.43, 0.53]
v_e	0.87 [0.85, 0.88]	0.66 [0.63, 0.70]	0.58 [0.54, 0.63]

*The brackets indicate 95% confidence intervals.

Table 3
Fraction of Variance Explained by Differences in Uterine Fibroids, Readers, and Random Error Using Random-Effect Analyses of Variance for Each Kinetic Parameter

	Large ROI	Targeted ROI	User-defined ROI
iAUC ($P < 0.0001$, $R^2 = 0.75$)			
Uterine fibroid	82.9%	73.0%	68.5%
Reader	3.0%	1.5%	10.3%
Error	14.1%	25.6%	21.2%
K^{trans} ($P < 0.0001$, $R^2 = 0.75$)			
Uterine fibroid	89.4%	74.1%	67.7%
Reader	1.6%	1.5%	14.1%
Error	9.0%	24.4%	18.2%
k_{ep} ($P < 0.0001$, $R^2 = 0.60$)			
Uterine fibroid	98.2%	46.9%	55.9%
Reader	0.1%	0.5%	0.4%
Error	1.7%	52.6%	43.7%
$\mathbf{v_e}$ ($P < 0.0001$, $R^2 = 0.69$)			
Uterine fibroid	88.7%	72.3%	64.8%
Reader	0.6%	<0.1%	4.2%
Error	10.7%	27.7%	31.0%

P values and adjusted R^2 values refer to model significance. The readers and uterine fibroids were considered random effects and the ROI method was considered a fixed effect.

input function (AIF), presence or absence of background T1 correction, region of interest (ROI) size, ROI placement, etc.). This becomes particularly challenging for multi-site studies. However, even if all study parameters are maintained, there is still the considerable possibility for intra- and inter-observer variation, as well as errors or variance in the postprocessing method. Our study is an attempt to quantify this variation in a rigorous manner for TWIST-derived DCE datasets using identical source data and postprocessing methods.

There exists previous work that has assessed DCE-MRI reproducibility. Padhani et al in 2001 (13) directly measured the intra-observer reproducibility of K^{trans} and v_e in normal tissue (muscle and bone) for patients with prostate cancer who were undergoing androgen deprivation therapy. They used single-section spoiled gradient echo fast low angle shot (FLASH) or five-section saturation recovery turbo-FLASH sequences to acquire DCE T1-weighted datasets with a temporal resolution of 9–10 s over 6.3–7.0 min. Two datasets using identical protocols were obtained within a mean time of 119 days. They found that the within-subject K^{trans} values had such substantial variation that it precluded CV calculations (mean K^{trans}: 0.126–0.554/min, SD: 0.189–0.331/min), while the v_e values had a within-subject CV of 19–26%. A similar study (14) using paired (obtained within one week) T1-weighted spoiled gradient echo FLASH datasets found similar magnitude within-subject CVs for a diverse set of malignant masses (K^{trans} 29%, k_{ep} 24%, v_e 9%, iAUC 12%), and greater magnitude within-subject CVs for benign muscle (k_{ep} 49%, v_e 16%, iAUC 17%). The within-subject K^{trans} CV for normal muscle could not be calculated. For the malignant masses in this study, the authors predicted that K^{trans} would need to decrease by 45% or increase by 83% within an individual to be confident that the change was not due to error or background fluctuation.

Other studies have reported within-subject variation for a variety of tissue types ranging from 8.9% to 24% for the quantitative kinetic parameters assessed in this study (15–20). These studies have also in general relied upon measurement comparisons obtained at two separate time points in the same patient without intervening therapy. Therefore, those measures of within-subject variation include any biologic variation that might occur from day-to-day in a given patient, variation induced by differences in patient positioning, as well as variation that might occur within or between MR systems. Our study eliminated that variation by simply assessing the within-subject variation in an identical DCE dataset, making measurements at multiple times in different ways within and between readers. Thus, our within-subject CVs (4.6–7.6%) were lower than published data, representing a lower-bound on variability given controlled circumstances. We would expect that if data were obtained using this technique in the same patients on different days, with more ill-defined masses, our CVs would be even greater (i.e., measurement error + biologic error).

In either case, our data strongly suggest that reader variation (within and between), ROI selection method, and possibly postprocessing variation (despite identical input parameters) are noteworthy impediments to quantitative DCE-MRI reproducibility. Although seemingly little can be done to improve intra- and inter-rater variation (other than adequate training), consistency in ROI selection method is important. In clinical practice, it may be difficult to replicate a perfusion measurement during longitudinal follow-up if specific details regarding the ROI selection method are not included in the Radiology report. Our data demonstrate that the user-directed ROI had the least inter-rater reproducibility. When readers were given the freedom to place the ROI where they believed most appropriate, the variation increased. In our study, the large ROI had the best inter-rater reproducibility. Larger ROIs can average differences in perfusion within a heterogeneous mass and are better able to accommodate differences between readers.

We intentionally selected uterine fibroids for our model because they were easy for readers to identify, circumscribed, and represented a range of perfused tissue (from centrally necrotic to homogeneously enhancing to nonenhancing). All of the uterine fibroids in this study were sharply demarcated and greater than 1.5 cm in diameter. We believed these features would reduce error introduced by difficulty with lesion identification and thereby reduce the influence of reader experience. Our data cannot be directly extrapolated to malignant tissue, particularly those with less conspicuous margins (in which case, we would hypothesize that the variability in our measurements would increase), because uterine fibroids are not malignant. However, our data do provide some insight into the variability that might still arise from multiple readers performing the same measurements on the same cases, postprocessed using the same workstation.

There are some limitations of our study. Our DCE datasets were derived using a TWIST sequence, which is different from the traditional 2D or 3D T1-weighted spoiled gradient echo sequences used in most DCE studies. We use TWIST to optimize the frame rate, but the unique method of *k*-space undersampling TWIST uses may limit our ability to generalize the results of this study to other DCE imaging protocols. Our measured variability could be more or less with a traditional protocol. Additionally, each of the five readers in this study was either a senior radiology resident or a board-certified abdominal imaging fellow who had received prestudy training on DCE analysis. It is possible that readers with more experience using the postprocessing tool or interpreting pelvic MR in general would have improved intra-rater reproducibility. However, each reader was trained and then evaluated by an experienced MR reader before study initiation to confirm training adequacy. Each reader was, therefore, skilled in the use of the postprocessing tool and competent with respect to all ROI placement methodologies before beginning the study.

In conclusion, under relatively constant conditions the mean intra-rater within-subject CVs for K^{trans}, k_{ep}, v_e, and iAUC using the same ROI selection method and postprocessing algorithm on identical TWIST DCE-MRI datasets were 4.6–7.6%, suggesting that changes of up to 9–15% may be attributable to measurement error. A large ROI encompassing at least three-quarters of the lesion had the greatest inter-rater reproducibility, while the user-defined ROI had the least inter-rater reproducibility. Our findings represent measurements performed in a uterine fibroid model and may not be a true reflection of general oncologic practice. Further studies would be needed to validate our results in that context.

REFERENCES

1. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusible tracer: standardized quantities and symbols. J Mag Reson Imaging 1999;10:223–232.
2. Koh TS, Thng CH, Hartono S, et al. A comparative study of dynamic contrast-enhanced MRI parameters as biomarkers for anti-angiogenic drug therapy. NMR Biomed 2011;24:1169–1180.
3. Knopp MV, Giesel FL, Marcos H, et al. Dynamic contrast-enhanced magnetic resonance imaging in oncology. Top Magn Reson Imaging 2001;12:301–308.
4. Walker-Samuel S, Leach MO, Collins DJ. Evaluation of response to treatment using DCE-MRI: the relationship between initial area under the gadolinium curve (IAUGC and quantitative pharmacokinetic analysis. Phys Med Biol 2006;51:3593–3602.
5. Padhani AR. MRI for assessing antivascular cancer treatments. Br J Radiol 2003;76:S60–S80.
6. Casneuf VF, Delrue L, Van Damme N, et al. Noninvasive monitoring of therapy-induced microvascular changes in a pancreatic cancer model using dynamic contrast-enhanced magnetic resonance imaging with P846, a new low-diffusible gadolinium-based contrast agent. 2011;175:10–20.
7. Preda A, van Vliet M, Krestin GP, et al. Magnetic resonance macromolecular agents for monitoring tumor microvessels and angiogenesis inhibition. Invest Radiol 2006;41:325–331.
8. Turetschek K, Preda A, Floyd E, et al. MRI monitoring of tumor response following angiogenesis inhibition in an experimental human breast cancer model. Eur J Nucl Med Mol Imaging 2003; 30:448–455.
9. Turkbey B, Thomasson D, Pang Y, et al. The role of dynamic contrast-enhanced MRI in cancer diagnosis and treatment. Diagn Interv Radiol 2010;16:186–192.
10. Homer J, Beevers MS. Driven-equilibrium single-pulse observation of T1 relaxation. A reevaluation of a rapid "new" method for determining NMR spin-lattice relaxation-times. J Magn Reson 1985;63:287–297.
11. Tofts PS, Kermode AG. 1991 Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. Magn Reson Med17:357–367.
12. Parker GJ, Roberts C, Macdonald A, et al. Experimentally-derived functional form for a population-averaged high temporal-resolution arterial input function for dynamic contrast-enhanced MRI. Magn Reson Med 2006;56:993–1000.
13. Padhani AR, Hayes C, Landau S, Leach MO. Reproducibility of quantitative dynamic MRI of normal human tissues. NMR Biomed 2002;15:143–153.
14. Galbraith SM, Lodge MA, Taylor NJ, et al. Reproducibility of dynamic contrast-enhanced MRI in human muscle and tumours: comparison of quantitative and semi-quantitative analysis. NMR Biomed 2002;15:132–142.
15. Ng CS, Raunig DL, Jackson EF, et al. Reproducibility of perfusion parameters in dynamic contrast-enhanced MRI of lung and liver tumors: effect on estimates of patient sample size in clinical trials and on individual patient responses. AJR Am J Roentgenol 2010; 194:W134–W140.
16. Lankester KJ, Taylor NJ, Stirling JJ, et al. Effects of platinum/taxane based chemotherapy on acute perfusion in human pelvic tumours measured by dynamic MRI. Br J Cancer 2005;93: 979–985.
17. Jackson A, Jayson GC, Li KL, et al. Reproducibility of quantitative dynamic contrast-enhanced MRI in newly presenting glioma. Br J Radiol 2003;76:153–162.
18. Ashton E, Raunig D, Ng C, et al. Scan-rescan variability in perfusion assessment of tumors in MRI using both model and data-derived arterial input functions. J Magn Reson Imaging 2008;28: 791–796.
19. Ferl GZ, Xu L, Friesenhahn M, et al. An automated method for nonparametric kinetic analysis of clinical DCE-MRI data: application to glioblastoma treated with bevacizumab. Magn Reson Med 2010;63:1366–1375.
20. Roberts C, Issa B, Stone A, et al. Comparative study into the robustness of compartmental modeling and model-free analysis in DCE-MRI studies. J Magn Reson Imaging 2006;23:554–563.