MRI signal is described by hardware-specific factors, proton density (PD) signal scaling factors, voxel volume, and pulse-sequence weighting, whereas quantitative MRI uses Bloch-Torrey equations that govern pulse sequence weighting components to distill tissue-specific properties such as T1, T2, and PD (1,2). Based on quantitative MRI data, synthetic MRI can generate both qualitative and quantitative images simultaneously from parental data (2–8).

Quantitative MRI mapping techniques of the knee allow the early characterization and quantification of articular abnormalities and effects of therapeutic interventions (9,10), whereas morphologic T1-weighted, intermediate-weighted, T2-weighted, and short-tau inversion recovery (STIR) MR images allow the characterization of structural abnormalities. However, the separate acquisitions of quantitative and qualitative images can be time consuming, and thus synthetic MRI may be advantageous by offering the simultaneous generation of quantitative maps and morphologic images. Synthetic MRI has been successfully used in the brain (11,12), but its role is less well established for MRI of the knee.

We tested the hypothesis that synthetic MRI of the knee generates accurate and repeatable quantitative maps and produces morphologic MR images with similar quality and detection rates of structural abnormalities similar to those of conventional MR images.
Synthetic MRI of the Knee

Abbreviations
CNR = contrast-to-noise ratio, ISMRM = International Society for Magnetic Resonance in Medicine, NIST = National Institute of Standards and Technology, PACS = picture archiving and communication system, PD = proton density, ROI = region of interest, SNR = signal-to-noise ratio, STIR = short-tau inversion recovery

Summary
Synthetic QRAPMASTER MRI of the knee is accurate for T1, T2, and proton density quantification, and simultaneously generated synthetic morphologic MR images have detection rates of structural abnormalities similar to those of conventional MR images, with similar acquisition time.

Implications for Patient Care
- Based on quantitative QRAPMASTER data, synthetic MRI of the knee generates quantitative maps and morphologic MR images with the same acquisition time required for conventional morphologic MRI.
- Synthetic MRI of the knee is accurate for T1, T2, and proton density quantification with phantom-based model-corrected average error margin of 0.8%.
- Synthetically generated morphologic MR images using the QRAPMASTER technique have higher contrast resolution of cartilage and meniscus relative to joint fluid when compared with conventional MRI, and similar detection rates for structural abnormalities as conventional MRI with similar acquisition time.

MRI Technique
We used a commercially available wide-bore 3-T MRI system (Magnetom Skyra, Numaris/4 Syngo MR E11C; Siemens Healthcare) with 48 independent radiofrequency receiver channels, maximum gradient field amplitude of 45 mT/m, and a slew rate of 200 T/m/sec. For phantom measurements, a head coil (Siemens Healthcare) with 16 receiver channels was used. For human participants, a knee coil (Quality Electrodynamics, Mayfield, Ohio) with one transmit and 15 receiver channels was used.

For the acquisition of parental synthetic MRI data, we used a sagittally oriented, biphasic QRAPMASTER (quantification of relaxation times and proton density by multiecho acquisition of a saturation-recovery using turbo spin-echo readout) MRI pulse sequence prototype (6,7). The pulse sequence used a two-dimensional, multisection, multiecho, multisaturation delay saturation-recovery turbo spin-echo technique with a repetition time of 4000 msec, two echo times of 21 and 103 msec, inversion time of 27 msec, four saturation delay times of 150, 580, 1860, and 3860 msec, parallel acceleration factor of 3, echo train length of five, receiver bandwidth of 401 Hz per pixel, flip angle of 150°, field of view of 160 × 160 mm², matrix of 320 × 240, section thickness of 3 mm and 0.3 mm intersection gap, and 28 sections (Table 1). The acquired data were used to generate quantitative T1, T2, and PD maps and quantitative T1-weighted, intermediate-weighted, T2-weighted, and STIR MR images by using commercially and publicly available software (SyMRI NEURO, version 8.0.4; SyntheticMR AB). The commercially and publicly available SyMRI NEURO software package was characterized by advanced functions, including the ability to export and transfer synthesized images in Digital Imaging and Communications in Medicine format to our picture archiving and communication system (PACS) (Vue version 12.1.0.2041; Carestream Health, Rochester, NY) for observer evaluations. Additionally, this software package afforded full synthetic functionality for synthesizing the entire spectrum of quantitative and morphologic musculoskeletal MR images and contrasts, without any restrictions to neuroradiological MRI.

For participants, we additionally acquired conventional T1-weighted, intermediate-weighted, T2-weighted, and STIR MR images with similar parameter settings (Table 1). The total acquisition times for conventional MRI and synthetic MRI for participants were 9 minutes 21 seconds and 9 minutes 52 seconds, respectively.

Phantom Evaluation
To validate the accuracy of the quantitative knee pulse sequence, we used an MRI system phantom developed by the International Society for Magnetic Resonance in Medicine (ISMRM) Ad Hoc Committee on Standards for Quantitative Magnetic Resonance and the National Institute of Standards and Technology (NIST) (13). The ISMRM-NIST phantom was considered the standard of reference, and synthetic MRI was considered the index test. The phantom consisted of T1, T2, and PD layers. Each layer contained 14 spheres with previously determined absolute T1 and T2 and PD percentage values at room temperature. We used spheres 1–6 in the T1 layer (351.5–1989 msec), spheres 1–10 in the T2 layer (22.56–581.3 msec), and spheres 1–14 in the PD layer (5%–100%). The MRI suite was set to 20°C. The bore fan was set on lowest convection. The phantom was given 12 hours to adapt to room temperature.

The ISMRM-NIST MRI phantom data acquisition was performed on 2 consecutive days to assess interday repeatability. On each day, two sessions were performed to assess intraday variability. During each session, each of the three layers was imaged twice at 30-minute intervals to assess intrainsession repeatability. In total, each layer of the phantom was imaged eight times. After each session, we repositioned the phantom in the coil and the coil in the MRI system. One observer (J.F.) with 15 years of musculoskeletal MRI experience performed measurements (SyMRI NEURO, version 8.0.4) of T1, T2, and PD values on synthetic T1, T2, and PD maps (Fig 1) using 1 cm² round regions of interest (ROIs). All measurements were repeated three times at 1-week intervals.

While the accuracy of the QRAPMASTER technique was assessed with the phantom measurements that were based on Bloch equations, heteroscedastic variation and residual errors of the quantitative data were then addressed through model-based correction by using logarithmic transformation and quadratic and split (segmented) quadratic equations. For PD data, which were expressed as percentage values and demonstrated no heteroscedasticity, logarithmic transformation was not required. Model-based correction was performed to reduce inhomogeneity.
Table 1: MRI Study Protocol

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Synthetic MRI, QRAPMASTER</th>
<th>Conventional MRI, Turbo Spin Echo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetition time (msec)</td>
<td>466 4000 4000 5860</td>
<td>466 4000 4000 5860</td>
</tr>
<tr>
<td>Echo time (msec)</td>
<td>7.9 31 102 30</td>
<td>7.9 31 102 30</td>
</tr>
<tr>
<td>Inversion time (msec)</td>
<td>— — 220 —</td>
<td>— — 220 —</td>
</tr>
<tr>
<td>Acceleration</td>
<td>3 3 3 3</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>Echo train length</td>
<td>5 5 5 5</td>
<td>3 15 15 17</td>
</tr>
<tr>
<td>Flip angle (degrees)</td>
<td>150 150 150 150</td>
<td>150 150 150 150</td>
</tr>
<tr>
<td>Receiver bandwidth (Hertz/pixel)</td>
<td>401 401 401 401</td>
<td>504 466 466 504</td>
</tr>
<tr>
<td>Field of view (mm)</td>
<td>160 × 160 160 × 160</td>
<td>160 × 160 160 × 160</td>
</tr>
<tr>
<td>Matrix</td>
<td>320 × 240 320 × 240</td>
<td>320 × 240 320 × 240</td>
</tr>
<tr>
<td>Section thickness/gap (mm)</td>
<td>3/0.3 3/0.3 3/0.3 3/0.3</td>
<td>3/0.3 3/0.3 3/0.3 3/0.3</td>
</tr>
<tr>
<td>No. of signals acquired</td>
<td>1 1 1 1 1</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>Concatenation</td>
<td>1 1 1 1 1</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>No. of sections</td>
<td>28 28 28 28</td>
<td>28 28 28 28</td>
</tr>
<tr>
<td>Phase-encoding direction</td>
<td>Anterior-to-posterior</td>
<td>Anterior-to-posterior</td>
</tr>
<tr>
<td>Acquisition time</td>
<td>...* ...* ...* ...*</td>
<td>1 min 54 sec 2 min 14 sec 2 min 16 sec 2 min 57 sec</td>
</tr>
</tbody>
</table>

Note.—IW = intermediate weighted, STIR = short-tau inversion recovery.

* For synthetic MRI (QRAPMASTER), the combined acquisition time for T1 weighted, IW, T2 weighted, and STIR images was 9 minutes 52 seconds.
of errors across parameter domains to maintain accuracy at the extremes of the included relaxation times. Heteroscedastic variation can occur due to additive Gaussian noise at longer repetition and echo times and monoeponential fitting not accounting for the effects of unmodeled variables such as spatially varying gradients, magnetization transfer, and anisotropy (8,14,15). Models were fit by using residual errors determined by ordinary least-squares, as the transformations induced reasonable homoscedasticity.

**Participant Evaluation**

Between January 2017 and April 2018, 54 participants (mean age, 40 years; age range, 18–62 years) including 24 men (mean age, 37 years; range, 18–62 years) and 30 women (mean age, 40 years; range, 21–60 years) were recruited from our practice (Fig 1). Indications for knee MRI were made in accordance with published guidelines (16). Each participant underwent our MRI study protocol twice on the same day (Table 1). Between the two acquisitions, participants rested for 30 minutes in a chair. After each acquisition, we repositioned the coil in the MRI system. Fifteen of 54 participants (28%) underwent the MRI protocol again after 5 days on average, with a range of 1 to 9 days. For this study part, conventional MRI was considered the standard of reference and synthetic MRI, the index test. Following data acquisition, the synthetic MRI data were exported to a network drive, imported into the dedicated software (SyMRI NEURO, version 8.0.4) where the quantitative maps and morphologic MR images were synthesized with a semiautomatic preset protocol on a standard desktop computer, and finally sent to our PACS (Vue version 12.1.0.2041; Carestream Health). This process required approximately 5 minutes or less.

Quantitative outcome variables included intraday and interday repeatability of T1, T2, and PD measurements of cartilage on quantitative maps, as well as signal-to-noise (SNR) and contrast-to-noise (CNR) ratios of fluid, cartilage, meniscus, marrow, and muscle on morphologic T1, intermediate-weighted, T2, and STIR images of synthetic and conventional MRI.

T1, T2, and PD value measurements were performed in central patella articular cartilage by one observer (J.E.) (17). When patellar cartilage thickness was insufficient, central trochlear cartilage was measured. The mean pixel value of oval 0.1 cm² ROIs sampling approximately 40 pixels was used (SyMRI NEURO, version 8.0.4). Measurements were repeated three times with 1-day intervals in between measurements. SNR and CNR were measured (SyMRI NEURO, version 8.0.4) in cancellous bone (distal femoral metaphysis), articular cartilage (patella or alternatively trochlear cartilage), joint fluid (intercondylar notch), and meniscus (posterior horn of the medial or lateral meniscus) by one observer (J.F.). Round or oval ROIs were copied into identical locations on synthetic and conventional MR images. The mean pixel value of the ROIs was used as the signal intensity, whereas the mean standard deviation of a background ROI placed just anterior to the skin surface over the patella was used as the noise. SNR was determined as signal intensity of tissue divided by standard deviation of tissue. Subsequently, CNR was calculated as $|\text{SNR}_{\text{tissue}_1} - \text{SNR}_{\text{tissue}_2}|$. Measurements were repeated three times at 1-day intervals.

All qualitative outcome variables were obtained by two fellowship-trained musculoskeletal radiologists (B.F. and N.K.), with 5 and 10 years of musculoskeletal MRI experience, respectively. Evaluations were performed independently on randomized data sets after removal of all clinical and personal information. Disagreements of structural integrity and side-to-side comparison assessments were resolved through a final consensus interpretation. Assessments were performed with a standardized, equidistant, five-point Likert scale, where a rating of 1 denoted “very bad” with complete obscuration of

![Flow diagram of participants through the study. PD = proton density.](image)

---

**Inclusion criteria:**
- 18 years of age and older
- Agreement to participate
- Clinical indication for MRI

**Potentially eligible participants (62)**

- Contraindication for MRI
- Inability to fit the MRI coil
- Metallic implants

**Eligible participants (59)**

- Pacemaker (1)
- Large knee size (1)
- Metallic implant (1)

**Excluded participants (3/62, 5%)**

- Inability to complete MRI (2)
- Motion degradation (1)
- Full-thickness patellofemoral cartilage loss (2)

**Participants with index test (54)**

- Intra-day repeatability of T1, PD, and T2 quantification
- Signal-to-Noise and Contrast-to-Noise quantification
- Image quality
- Visibility of anatomical structures
- Integrity of anatomical structures
- Side-to-side comparison of conventional and synthetic MRI

**Participants with repeat index test (15/54, 28%)**

- Inter-day repeatability of T1, PD, and T2 quantification

---
anatomic details, and a rating of 5 denoted “very good” with the unimpaired depiction of all anatomic details. Assessments were performed on PACS software (Vue version 12.1.0.2041; Carestream Health). A 4 × 2 view-port setup was used with synchronized scrolling, sizing, and panning.

Image quality assessments included the degree of motion, noise, artifacts including chemical shift, interface and reconstruction artifacts, edge sharpness of structures, partial volume effects, contrast resolution defined as visual gray-scale differences between structures, fluid brightness, and fat suppression. An interface artifact manifests as an artificial linear signal along interfaces of different tissues and may occur if there is motion during acquisition.

Visibility of menisci, articular cartilage, cruciate ligaments, extensor tendons, and bone was evaluated in the context of internal derangement assessment on synthetic and conventional data sets consisting of T1-weighted, intermediate-weighted, T2-weighted, and STIR images.

The integrity of menisci, articular cartilage, anterior cruciate ligament, and subchondral bone was assessed. Meniscal tears were defined as substance defect extending to the articular surface. Articular cartilage defects were defined as substance loss greater than 50%. Only the largest articular cartilage defect was assessed. Anterior cruciate ligament tears were defined as 50% or greater substance loss of cross-sectional area. Bone marrow edema was defined as STIR signal hyperintensity compared with the distant normal marrow. Discrepant findings were resolved during consensus interpretation.

Both observers performed a side-to-side comparison, rating corresponding synthetic and conventional T1-weighted, intermediate-weighted, T2-weighted, and STIR images as superior, inferior, or equal based on their subjective impression of the suitability of the images for accomplishing an evaluation for internal knee derangement.

Statistical and Quantitative Assessment
Statistical analyses were performed by using R 3.3 software with lme4 and epiR packages (http://cran.r-project.org). Variables are given as the average with standard deviation, median with minimum and maximum in parentheses, ratios, or percentages. For the evaluation of the quantitative outcome variables, an apriori Wilcoxon signed-rank test for related samples power calculation derived a minimum sample size of 26 participants, an apriori Wilcoxon signed-rank test for related samples or xferences in the comparison assessments in participants were assessed with the Wilcoxon test for related samples or χ² test. The coefficient of variation was used to assess the precision of measurements. The interobserver and intermethod agreements in participants were determined by using the Cohen kappa test with linear weights for Likert scale assessments and the Cohen kappa test without weights for binary assessments. Kappa values were graded according to Landis and Koch (18). In the case of acceptable agreement, observer assessments were combined. P values less than or equal to .01 were considered to indicate a statistically significant difference.

Results

Phantom Evaluation
The native relative errors of measured T1, T2, and PD values compared with phantom reference values, adjusted for day, session, and replicate variations, were 1.9%, 7.4%, and 5.1%, respectively, whereas the model-corrected relative errors of measured T1, T2, and PD values were 0.8%, 1.4%, and 0.3%, respectively. The average relative error of measured values to reference values was 0.5% (0.1%–0.9%) for intrasession measurements, 0.8% (0.6%–1.2%) for intersession measurements, and 1.0% (0.7%–1.1%) for interday measurements. The average overall relative error of measured values compared with reference phantom values was 0.8% (0.3%–1.4%) following model correction.

The T1 measurements demonstrated a heteroscedastic variation. Fitting of log-linear and basic quadratic nonlinear models yielded the following calibration equation with an accuracy of 0.8% (Fig 2):

\[
\text{Reference [msec]} = -36.09 + \sqrt{1302.42 + 1.59 \times \text{Measurement [msec]}^{1.95}}
\]

The T2 measurements demonstrated a heteroscedastic variation as well. Fitting of a split quadratic model with a B-splines approach yielded the following calibration equation with an accuracy of 1.4% (Fig 2):

\[
\begin{align*}
\text{Reference [msec]} &= -43.53 + \sqrt{1895.19 + 11.10 \times \text{Measurement [msec]}^{0.62}} \\
&\quad \text{if Measurement < 195.63 msec} \\
&= -935.03 + \sqrt{1020698 + 51.72 \times \text{Measurement [msec]}^{0.62}} \\
&\quad \text{otherwise}
\end{align*}
\]

The PD measurements demonstrated homoscedastic variation. Fitting of a split quadratic model with a B-splines approach yielded the following calibration equation with an accuracy of 0.3% (Fig 2):

\[
\begin{align*}
\text{Reference [pu]} &= 97.13 - \sqrt{9433.65 - 195.82 \times \text{Measurement [pu]}^{0.01}} \\
&\quad \text{if Measurement < 32.95 pu} \\
&= -24.92 + \sqrt{-1504.66 + 145.95 \times \text{Measurement [pu]}^{0.01}} \\
&\quad \text{otherwise} \\
&= 164.02 - \sqrt{19994.57 - 155.14 \times \text{Measurement [pu]}^{0.01}} \\
&\quad \text{if Measurement > 66.30 pu}
\end{align*}
\]

Participant Evaluation
Intraday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference
Figure 2: Graphs depict measurement accuracy of synthetic quantitative T1, T2, and proton density (PD) images with an International Society for Magnetic Resonance in Medicine—National Institute of Standards and Technology MRI phantom.

Figure 3: Bland-Altman plots of in vivo intraday and interday agreement of quantitative T1, T2, and proton density (PD) measurements on synthetic MRI quantitative maps. Each symbol represents one participant. Three symbols of the same kind represent three repeat measurements.
Table 2: In Vivo Repeatability of Raw and Model-corrected Quantitative Synthetic MRI Data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Native Synthetic MRI Data</th>
<th>Model-corrected Synthetic MRI Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Difference (%)</td>
<td>Maximum Difference (%)</td>
</tr>
<tr>
<td>T1</td>
<td>3.4 ± 2.0</td>
<td>3.8 ± 3.1</td>
</tr>
<tr>
<td>PD</td>
<td>3.2 ± 2.3</td>
<td>3.8 ± 3.1</td>
</tr>
<tr>
<td>T2</td>
<td>3.3 ± 2.4</td>
<td>4.4 ± 3.8</td>
</tr>
<tr>
<td>All</td>
<td>3.3 ± 2.4</td>
<td>3.5 ± 2.9</td>
</tr>
</tbody>
</table>

Note.—PD = proton density, CV = coefficient of variation.

* Data are mean ± standard deviation.

of 4.1% (minimum, 0.1%; maximum, 12.4%) (Fig 3). The coefficient of variation of measurements was 1.1% for both the first and second session.

Interday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference of 3.3% (0.3%–9.4%) (Fig 3). The coefficient of variation of measurements was 1.2% for both the 1st and 2nd days.

After model correction with phantom data–derived equations, the interday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference of 3.5% (0.3%–9.6%) (Table 2). The coefficient of variation of model-corrected measurements was 1.3% for both the 1st and 2nd days. The average repeatability coefficient was 21.86 (6.8%).

SNR and CNR ratios of different tissues of morphologic synthetic and conventional T1-weighted, intermediate-weighted, T2-weighted, and STIR MR images and their comparison are given in Figure 4. On synthetic T1-weighted images, SNR of fluid was lower ($P < .001$). On synthetic intermediate-weighted and T2-weighted MR images, SNR of cartilage and SNR of fluid was higher ($P < .001$, respectively). On synthetic STIR images, SNR of fluid was higher ($P < .001$) and SNR of bone marrow and SNR of menisci was lower ($P < .001$, respectively). On synthetic T1-weighted images, the fluid-to-menisci CNR was lower ($P < .001$) and cartilage-to-fluid CNR was higher ($P < .001$). On synthetic intermediate-weighted, T2-weighted, and STIR images, the cartilage-to-fluid CNR, menisci-to-fluid CNR, and muscle-to-fluid CNR was higher ($P < .001$).

Image quality assessments (Table 3) showed synthetic MRI had greater STIR fat suppression ($P < .001$) and fluid signal ($P = .10$), as well as higher degrees of image noise ($P = .001$) and artifacts ($P < .001$) (Fig 5). There were no differences between the other image quality parameters (Table 3).

Visibility of menisci, articular cartilage, anterior and posterior cruciate ligaments, extensor tendons, and bone was rated as good to very good on conventional and synthetic STIR, T1-, intermediate-, and T2-weighted MR images, with interobserver agreements ranging from moderate to good (kappa, 0.584–0.708) and no differences noted ($P$ values = .01–.73). Table 4 shows the frequencies of meniscal tears (Fig 6), articular cartilage defects (Fig 5), and areas of bone marrow edema. There were no anterior cruciate ligament and extensor mechanism tears. The interobserver agreements were moderate to very good. The intermethod agreements were good. Among 108 potential discrepancies between conventional and synthetic MRI for both observers of each structure, there were 11 (10%) for medial meniscus, nine (8%) for lateral meniscus, 22 (20%) for articular cartilage defects, and 11 (10%) for bone marrow edema.

For side-to-side comparison, observer A rated synthetic MRI in six of 54 (11%) and conventional MRI in three of 54 (6%) participants as superior, whereas 45 of 54 (83%) were rated as equivalent. Observer B rated synthetic MRI in three of 54 (6%) and conventional MRI in six of 54 (11%) participants as superior, whereas 45 of 54 (83%) were rated as equivalent ($\chi^2 = 16$, $P = .003$).
Discussion

We report the native and model-corrected accuracy of synthetic knee MRI for T1, T2, and PD quantification using an ISMRM-NIST phantom and show high intraday and interday repeatability in living human participants. All synthetic MR images showed improved CNR for cartilage evaluation, and synthetic T2-weighted, intermediate-weighted, and STIR MR images showed improved CNR for meniscal evaluation. Observers perceived improvement of STIR fat suppression with synthetic MRI, whereas the overall quality ratings and detection rates of various internal knee derangements were similar with synthetic and conventional MRI.

The validation of the accuracy of synthetic MRI against a standard of reference is a prerequisite for its clinical use and appropriate patient care. Therefore, we validated and model-corrected the QRAPMASTER technique against an internationally accepted quantitative MRI phantom (13). Our approach contrasts attempts of validation that compared T2 relaxation times with other quantitative fast-spin-echo multiecho techniques (19), which introduce inaccuracies related to monoexponential T2 curve fitting (20,21) and, therefore, may not be representative of conventional single echo time fast-spin-echo T2-weighted techniques. Multiecho methods may also produce tissue-specific T2 relaxation differences when compared with conventional sequences and phantom-validated disagreements at echo times of less than 19 msec (20,22). A prior, uncalibrated phantom evaluation of a synthetic MRI prototype technique with four echo times and limited coverage of the T1, T2, and PD spectra showed T1 and T2 relaxation time underestimation and PD percentage overestimation of $-1.2\% \pm 5.4$, $-6.6\% \pm 1.5$, and $0.8\% \pm 1.5$, respectively (23). In comparison, our phantom experiment sampled larger T1 (351–1989 msec), T2 (22–581 msec), and PD (5%–100%) domains, which better encompass the physiologic range of structures of the knee (23–25).

We show that native accuracy of the quantitative data varies in a heteroscedastic manner and that model correction
should be performed for overall reduction and error homogenization across parameter domains to maintain accuracy at the extremes of the relaxation rate curves. Such correction reduced the relative accuracy errors of measured T1, T2, and PD values from 1.9%, 7.4%, and 5.1% to 0.8%, 1.4%, and 0.3%, respectively. Our average phantom-based mean-adjusted percentage accuracy errors compare favorably with phantom-based accuracy errors of mixed-echo turbo-spin-echo techniques of 1.6%–10.9% for T1 and 9.4%–12.9% for T2 (4), 10%–13% for T2 with a multicomponent quantitative technique (26), 0%–8.3% for T1 with modified Look-Locker technique, 0%–1.2% for T1 with saturation.

**Table 3: Observer Ratings of Image Quality Parameters of Synthetic and Conventional MRI Methods**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Conventional MRI*</th>
<th>Synthetic MRI*</th>
<th>Intermethod P Value</th>
<th>Interobserver Agreement†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion</td>
<td>5 (3, 4–5, 5)</td>
<td>5 (3, 4–5, 5)</td>
<td>.50</td>
<td>0.591 (0.435, 0.747)</td>
</tr>
<tr>
<td>Noise</td>
<td>5 (4, 4–5, 5)</td>
<td>4 (3, 4–5, 5)</td>
<td>.001</td>
<td>0.63 (0.483, 0.776)</td>
</tr>
<tr>
<td>Artifact</td>
<td>5 (4, 4–5, 5)</td>
<td>4 (3, 3–4, 4)</td>
<td>&lt;.001</td>
<td>0.432 (0.282, 0.582)</td>
</tr>
<tr>
<td>Edge sharpness</td>
<td>4 (3, 4–5, 5)</td>
<td>4 (3, 4–5, 5)</td>
<td>.63</td>
<td>0.668 (0.532, 0.804)</td>
</tr>
<tr>
<td>Partial volume effects</td>
<td>4 (3, 3–5, 5)</td>
<td>4 (3, 3–5, 5)</td>
<td>.76</td>
<td>0.689 (0.573, 0.805)</td>
</tr>
<tr>
<td>Contrast resolution</td>
<td>4 (4, 4–5, 5)</td>
<td>4 (4, 4–5, 5)</td>
<td>.90</td>
<td>0.631 (0.482, 0.781)</td>
</tr>
<tr>
<td>Fluid signal</td>
<td>4 (4, 4–5, 5)</td>
<td>5 (4, 4–5, 5)</td>
<td>.10</td>
<td>0.463 (0.295, 0.63)</td>
</tr>
<tr>
<td>Fat suppression</td>
<td>4 (3, 3–4, 4)</td>
<td>5 (3, 4–5, 5)</td>
<td>&lt;.001</td>
<td>0.57 (0.435, 0.706)</td>
</tr>
</tbody>
</table>

* Based on a five-point Likert scale, where 1 is the lowest value (“very bad”) and 5 the highest value (“very good”).
† Data are κ values, with 95% confidence intervals in parentheses.
recovery single shot acquisition technique, and 5%–15% for T2 with steady-state free precession technique (8).

We applied phantom-derived model corrections to living participants to improve in vivo accuracy. Since nonlinear model correction may unpredictably affect repeatability, we demonstrate near equivalent in vivo repeatability using split-quadratic model corrections of logarithmized data that account for heteroscedasticity and residual error structure. The phantom-based accuracy and subject-based precision errors of T1, T2, and PD quantification appear at least acceptable for clinical use. A study investigating the Osteoarthritis Initiative cohort demonstrated a significant increase of cartilage T2 relaxation times over a period of 6 years, from 32 msec to 34 msec (6.3%) in participants with simultaneous worsening in the whole organ MRI cartilage score (27). Our model-corrected T2 phantom-based accuracy error of 1.4% and subject-based precision error of 4.4% suggest the capability of our technique for detecting such a magnitude of change, which may contrast with previously reported T2 accuracy errors of 5%–15% with steady-state free precession technique (8) and 10%–13% with a Carr-Purcell-Meiboom-Grill pulse sequence (26). In addition, our corrected T1 repeatability error of 3.6% (range, 0.4%–10.1%) compares favorably to a prior conventional MRI phantom multicenter study of variable-flip-angle T1 quantification (28), which found a repeatability median error range of 0.7%–25.8% for T1 quantification.

Our initial results suggest similar detection rates with synthetic and conventional MRI for structural abnormalities of the knee; however, larger studies and correlation with arthroscopic surgery are needed to define diagnostic accuracies. Improved CNR between cartilage and fluid and menisci and fluid on synthetic T2-weighted, intermediate-weighted, and STIR MR images may help to diagnose subtle abnormalities. Synthetic MR images had a small, but higher degree of interface artifacts, which may interfere with the detection of subtle signal abnormalities at the tidemark of articular cartilage. We noticed improved STIR fat suppression with QRAPMASTER, which we believe is in part the result of B1 inhomogeneity correction.
Table 4: Observer Assessments of Internal Derangement with Conventional and Synthetic MRI Methods

<table>
<thead>
<tr>
<th>Structural Abnormality</th>
<th>Conventional MRI</th>
<th>Synthetic MRI</th>
<th>Intermethod Agreement‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial meniscus tear</td>
<td>14 (26)</td>
<td>13 (24)</td>
<td>Observer 1 Observer 2 Consensus</td>
</tr>
<tr>
<td></td>
<td>(0.627, 0.988)</td>
<td>(0.524, 0.955)</td>
<td>0.798 (0.609, 0.957)</td>
</tr>
<tr>
<td>Lateral meniscus tear</td>
<td>13 (24)</td>
<td>12 (22)</td>
<td>Observer 1 Observer 2 Consensus</td>
</tr>
<tr>
<td></td>
<td>(0.524, 0.955)</td>
<td>(0.524, 0.955)</td>
<td>0.74 (0.422, 0.989)</td>
</tr>
<tr>
<td>Articular cartilage defect</td>
<td>16 (30)</td>
<td>17 (31)</td>
<td>Observer 1 Observer 2 Consensus</td>
</tr>
<tr>
<td></td>
<td>(0.524, 0.955)</td>
<td>(0.524, 0.955)</td>
<td>0.798 (0.609, 0.957)</td>
</tr>
<tr>
<td>Bone marrow edema</td>
<td>22 (41)</td>
<td>20 (37)</td>
<td>Observer 1 Observer 2 Consensus</td>
</tr>
<tr>
<td></td>
<td>(0.524, 0.955)</td>
<td>(0.524, 0.955)</td>
<td>0.74 (0.422, 0.989)</td>
</tr>
<tr>
<td>All</td>
<td>66 (122)</td>
<td>75 (139)</td>
<td>0.721 (0.561, 0.944)</td>
</tr>
</tbody>
</table>

* No anterior cruciate ligament tears were seen.
† Data in parentheses are percentages.
‡ Data are κ values, with 95% confidence intervals in parentheses.

With use of local effective flip angles (21). B1 inhomogeneity correction may also result in improved T1 contrast and account for our observation that bone marrow edema is particularly hypointense on synthetic T1-weighted MR images. As T1 hypointensity of bone marrow lesions relative to muscle is a frequently used imaging sign for marrow replacement (29), synthetic MR images may paradoxically decrease the specificity of this criterion and require additional chemical shift imaging or fat-fraction quantification for definitive evaluation. Given this finding, there is also the potential for synthetic MRI to correct for T1 bias in the fat-fraction quantification of bone marrow abnormalities without lowering flip angles, which reduces the SNR (30).

The efficiency of synthetic MRI in a clinical setting may depend on whether the total acquisition time is less than that with separately acquired conventional quantitative and morphologic MR images. In our study, synthetic and conventional MRI pulse sequence acquisition times differed by a few seconds; however, the QRAPMASTER sequence provides quantitative mapping as well as morphologic MR images in the same time that conventional MRI provides only morphologic MR images (19). While T1 mapping is most commonly used in conjunction with gadolinium-based contrast agents, T2 mapping may be the most frequently used non–gadolinium-based contrast agent technique for the detection and quantification of early cartilage degeneration. PD mapping is a promising parameter due to its association with histologic and biomechanical cartilage abnormalities (25), which can be obtained simultaneously with synthetic T2 maps. An additional potential benefit of synthetic MRI is the ability to simultaneously generate double inversion recovery images, such as STIR fluid-attenuated inversion recovery (FLAIR) images, which have been previously suggested as a replacement for postcontrast sequences in evaluating synovitis (31). However, we did not evaluate synthetic STIR FLAIR images in our study because intravenous contrast agent administration was not part of our study protocol.

Our study has limitations. We did not perform a conventional MRI comparison for the phantom experiment and did not test the derived model-correction equations in a second phantom or MR unit. Therefore, the unit- or phantom-specific systematic errors that may cause over- or undercorrection of QRAPMASTER data are unknown. However, our goal was not to produce generalizable model-correction equations, but instead to determine the in vitro accuracy improvement of the synthetic knee MRI pulse sequence with individual unit model corrections of T1, T2, and PD data and demonstrate maintained repeatability with the in vivo application of the model corrections. Additionally, our model-corrected phantom accuracies are congruent with prior synthetic and conventional MRI studies (20–22). The number of replications at each reference level, small variation associated with the replicates, and avoidance of complexity in curve fitting minimize over-fitting errors and make a training and testing set approach unnecessary. The similar proportions of internal derangement diagnosed by both observers with conventional and synthetic MRI suggest similar accuracies; however, agreements with surgical inspection are unknown. Owing to the size of the ISMRM-NIST phantom, we used a head coil for the phantom validation instead of the knee.
coil. While coil sensitivity is a static measure that T1 and T2 curve fitting compensate for (7), variations in knee position and differences of knee morphology may have contributed to lower accuracy in humans. While the heteroscedastic error calibration is a function of the measured values and therefore applicable at human body temperature, correction for residual, temperature-related, substrate-dependent errors was not possible, which may have affected the in vivo accuracy, but not repeatability and detection of structural abnormalities.

In summary, synthetic QRAPMASTER MRI of the knee is accurate for T1, T2, and PD quantification and simultaneously generates morphologic MR images with high image contrast of cartilage and meniscus relative to joint fluid and similar detection rates of structural abnormalities when compared with conventional MRI with similar acquisition time.

Acknowledgments: We thank Martin Uppman, MSc, and Tobias Granberg, MD, PhD (Karolinska University Hospital, Stockholm, Sweden), and Frederik Testud, PhD (Siemens Healthcare AB, Sweden), for their work on the QRAPMASTER pulse sequence.