Radiotherapy and Oncology 133 (2019) 156–162

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Orignal Article

Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator

Ernst S. Kooreman [a], Petra J. van Houdt [a], Marlies E. Nowee [a], Vivian W.J. van Pelt [a], Rob H.N. Tijssen [b], Eric S. Paulson [c], Oliver J. Gurney-Champion [d], Jihong Wang [e], Folkert Koetsveld [a], Laurens D. van Buuren [a], Leon C. ter Beek [a], Uulke A. van der Heide [a,*]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [b] Department of Radiotherapy, University Medical Center Utrecht; [c] Department of Radiation Oncology, Medical College of Wisconsin, Milwaukee, United States; [d] Joint Department of Physics, The Institute of Cancer Research, and The Royal Marsden NHS Foundation Trust, London, United Kingdom; [e] Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, United States

ARTICLE INFO

Article history:
Received 5 September 2018
Received in revised form 4 January 2019
Accepted 9 January 2019
Available online 28 January 2019

Keywords:
Functional MRI
MR-linac
Multicenter
Quantitative MRI
Phantom
Prostate cancer

ABSTRACT

Purpose: Systems for magnetic resonance (MR-) guided radiotherapy enable daily MR imaging of cancer patients during treatment, which is of interest for treatment response monitoring and biomarker discovery using quantitative MRI (qMRI). Here, the performance of a 1.5 T MR-linac regarding qMRI was assessed on phantoms. Additionally, we show the feasibility of qMRI in a prostate cancer patient on this system for the first time.

Materials and methods: Four 1.5 T MR-linac systems from four institutes were included in this study. T_1 and T_2 relaxation times, and apparent diffusion coefficient (ADC) maps, as well as dynamic contrast enhanced (DCE) images were acquired. Bland–Altman statistics were used, and accuracy, repeatability, and reproducibility were determined.

Results: Median accuracy for T_1 ranged over the four systems from 2.7 to 14.3%, for T_2 from 10.4 to 14.1%, and for ADC from 1.9 to 2.7%. For DCE images, the accuracy ranged from 12.8 to 35.8% for a gadolinium concentration of 0.5 mM and deteriorated for higher concentrations. Median short-term repeatability for T_1 ranged from 0.6 to 5.1%, for T_2 from 0.4 to 1.2%, and for ADC from 1.3 to 2.2%. DCE acquisitions showed a coefficient of variation of 0.1–0.6% in the signal intensity. Long-term repeatability was 1.8% for T_1, 1.4% for T_2, 1.7% for ADC, and 17.9% for DCE. Reproducibility was 11.2% for T_1, 2.9% for T_2, 2.2% for ADC, and 18.4% for DCE.

Conclusion: These results indicate that qMRI on the Unity MR-linac is feasible, accurate, and repeatable which is promising for treatment response monitoring and treatment plan adaptation based on daily qMRI.

Biomarkers derived from quantitative magnetic resonance imaging (qMRI) are promising for oncology, where they provide functional information for treatment response monitoring and prediction. Dynamic contrast-enhanced (DCE-) MRI, which measures perfusion and permeability, and diffusion weighted imaging (DWI) MRI have shown to provide valuable information for different types of cancer [1,2]. Moreover, these modalities are increasingly recommended in guidelines for tumor staging and treatment response monitoring [3,4].

The use of qMRI biomarkers in clinical practice is currently limited [5]. Differences between systems, sequences, image reconstruction algorithms, and data processing methods all influence the results of biomarker studies and complicate the comparison of studies from different centers [6,7]. Moreover, studies investigating the use of qMRI biomarkers as surrogate endpoints have been limited to small patient groups. Larger cohort studies relating qMRI biomarkers to clinical outcome are yet to be done.

An additional challenge for treatment response monitoring, is that imaging studies often are performed at different time points [8–11], and evidence about the optimal timing is lacking. To our knowledge, only one study managed to accomplish daily MRI of patients during radiation treatment of brain metastases [12]. Such studies are challenging to perform, both logistically and in terms of patient burden, but are necessary to acquire more information about MRI related changes due to treatment response.

The recently introduced linear accelerators that are integrated with an MRI scanner (MR-linac), create the possibility of daily imaging during treatment with limited increase of patient burden

* Corresponding author at: Department of Radiation Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
E-mail address: u.vd.heide@nki.nl (U.A. van der Heide).

https://doi.org/10.1016/j.radonc.2019.01.011
0167-8140/

[13]. Treatments commonly take place over the course of weeks, where patients are daily positioned inside the MR-linac. This makes frequent MRI for treatment response monitoring and biomarker studies feasible. However, the design of the MRI part of an MR-linac is different from conventional diagnostic systems [14–18], which may influence the quality of the qMRI data. Therefore, before the start of treatment response monitoring and biomarker studies on an MR-linac, a thorough performance assessment is needed. Feasibility of qMRI has been shown previously on a hybrid MR-radiation therapy system that uses a combination of Cobalt-60 sources and a 0.35 T MRI, both qualitatively [19] as quantitatively using DWI [20,21]. These kinds of assessments are important in order to evaluate newly introduced MR-linac systems.

As a first step, this study aims to determine the accuracy, repeatability, and reproducibility of T_1 mapping, T_2 mapping, DWI, and DCE-MRI in phantoms on the Unity MR-linac (Elekta AB, Stockholm, Sweden), which is equipped with a modified Philips 1.5T MRI system (Philips Healthcare, Best, The Netherlands) [14]. We also show the feasibility of qMRI *in vivo* by acquiring quantitative maps in a prostate cancer patient.

Methods and materials

Study setup

Four qMRI acquisitions were performed: T_1 mapping, T_2 mapping, DWI for apparent diffusion coefficient (ADC) mapping, and DCE-MRI. We used the recommendations established by the quantitative imaging biomarker alliance (QIBA) DWI- and DCE-MRI profiles [22,23] for assessment of the performance of DWI and DCE-MRI data. For T_1 mapping and T_2 mapping we applied similar approaches as explained in detail below.

Phantom measurements were performed on four Unity MR-linac systems across four institutes. These systems are designated here as MR-linac A, B, C, and D. Each institute used its own copy of the phantoms described below, except institutes B and D, which used the same phantom for DWI.

On all systems, accuracy and short-term repeatability was assessed. The measurements were repeated within the scanning sessions for assesment of the short-term repeatability. To assess long-term repeatability, the measurements were repeated on system A after five months. Reproducibility among the four systems was determined as well. All image analysis and curve fitting were done using MATLAB (Release 2017b, MathWorks, Natick, MA).

In a prospective feasibility study, all qMRI acquisitions were performed on one patient with histologically proven prostate cancer. All protocols were approved by the medical ethics committee of The Netherlands Cancer Institute, and written informed consent was obtained.

T_1 mapping

The Eurospin TO5 phantom (Diagnostic Sonar, Livingston, Scotland) was used to evaluate the performance of T_1 mapping (see Fig. 1a). Twelve gel samples were chosen with T_1 relaxation times between 329 and 1603 ms at 296 K at 1.5 T. The variable flip angle (VFA) method was applied [24], using a spoiled gradient echo sequence with flip angles of 3, 6, 10, 20, and 30°. This method was chosen because it provides a fast way to map a 3D volume and is therefore often used clinically. The remaining acquisition parameters can be found in Table S1.

T_1 values were estimated from the mean values of regions of interest (ROIs) with a diameter of 12 mm per tube by a linear least-squares method [25]. Two institutes switched one tube with a tube of the DCE phantom for baseline T_1 measurements, so this tube was omitted in the analysis for all institutes. The temperature was measured before and after each acquisition in a tube with water that was kept near the phantom during all experiments. The average temperature was used to correct estimated T_1 values to the reference value at 296 K. For each tube, this reference value was provided by the phantom manufacturer at a field strength of 1.5 T and a temperature of 296 K.

T_2 mapping

The Eurospin TO5 phantom was used for T_2 mapping as well, but with a different set of gel samples (see Fig. 1b). The T_2 values ranged between 49 and 212 ms at 296 K at 1.5 T.

For acquisition, an accelerated multi-echo spin echo sequence was used. Acquisition parameters differed slightly for each system and can be found in Table S2.

Average decay curves were constructed for ROIs with a diameter of 12 mm within each tube. The T_2 values were estimated by fitting a mono-exponential decay function with a nonlinear least-squares method. To avoid bias from stimulated echoes, the first echo was discarded for analysis which was achieved by skipping the acquisition of the first echo during scanning [26]. As with the T_1 mapping, the temperature was monitored before and after each experiment, and the average temperature was used to correct estimated T_2 values to a reference value at 296 K.

DWI

A diffusion phantom (High Precision Devices, Inc, Boulder, Colorado) recommended by the QIBA DWI profile was used for DWI measurements (see Fig. 1c) [22]. This phantom contains 13 separate vials with aqueous solutions of 0–50% w/w polyvinylpyrrolidone. The vials were surrounded by ice water to ensure measurements at 0 °C. The phantom was placed such that the central tube was in the iso-center of the system. Each institute scanned the QIBA rec-

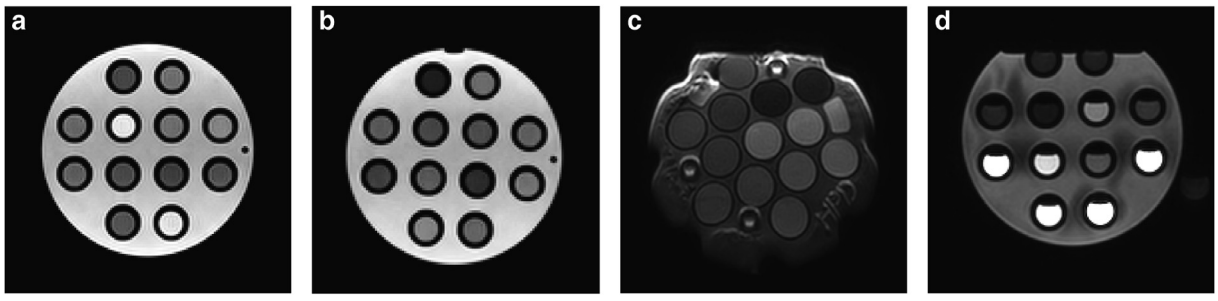

Fig. 1. MR images of the T1, T2, ADC, and DCE phantoms. (a) T1-weighted VFA image with flip angle of 10 degrees of the Eurospin TO5. (b) T2-weighted image with a TE of 70 ms. (c) DWI image ($b = 0$) of the QIBA phantom at the isocenter. (d) T1-weighted image of the DCE phantom.

ommended calibration protocol, which uses a spin-echo echo-planar imaging acquisition. Diffusion weighing was achieved using Stejskal-Tanner diffusion gradients with four b-values: 0, 500, 900, and 2000 s/mm^2. Other acquisition parameters can be found in Table S3. Additionally, a clinical protocol with larger voxels and b-values of 0, 200, and 800 was scanned on systems A and D (Table S7). Apparent diffusion coefficient (ADC) maps were calculated offline by fitting a linear function to the log of the signal decay versus *b*-values using linear least-squares. Mean values of the tubes were determined in a ROI with a diameter of 13 mm and compared to values provided by the phantom manufacturer.

DCE

The QIBA DCE profile [23] proposes to test three aspects of the acquisition: the accuracy of T_1 estimation, the stability of the signal during acquisition, and the linearity between signal intensity and concentration of the contrast agent. The first aspect was tested with the T_1 mapping of the Eurospin TO5 phantom. For the latter two, a phantom was created consisting of ten tubes with different concentrations of gadolinium between 0 and 9.8 mM (Dotarem, Guerbet, France, T_1 relaxivity 3.9 s^{-1} mM^{-1}) dissolved in a stock solution of water and 0.045 mM manganese chloride. These tubes were inserted in the Eurospin TO5 holder for image acquisition (see Fig. 1d).

A spoiled gradient-echo sequence was repeated for 4:39–5:17 minutes with a temporal resolution of 4.1–4.7 s, for a total of 65–75 scans. Additional acquisition parameters can be found in Table S4. Both the stability of the signal over all scans, and the relation between signal intensity and concentration was assessed. The signal intensity was measured as the mean of a ROI with a diameter of 5 mm in each tube. To ensure a steady state, the first two dynamics were discarded. The median value of the remaining dynamics was used for analysis.

Although the QIBA DCE profile suggests to assess the linearity of the signal intensities over the tubes with a range of contrast agent concentrations, we converted the signal intensities to concentration values to be able to compare the results between systems [27]. For this, a baseline T_1 value is needed, which was represented by the tube with 0 mM gadolinium. The T_1 value of this tube was determined separately using an inversion recovery (IR) series with inversion times of 30–4000 ms on each MR-linac individually. Acquisition parameters of the IR series are given in Table S5. The IR method was used because it is regarded the gold standard [28,29], so the influence of possible inaccuracies of the clinical T_1 mapping method was minimized allowing for a better assessment of the spoiled gradient-echo sequence. The calculated concentration values were compared to the known gadolinium concentrations.

Patient data

All quantitative measurements were obtained *in vivo* in a single patient, with similar settings as described above. Details on the sequence parameters are in Table S6. T_2- and ADC maps were calculated on the system. For DCE imaging, a T_1 map was estimated offline based on a VFA series [27]. The Tofts model was then applied to estimate K^{trans} [30], using an arterial input function with parameters derived from an in-house study population of prostate cancer patients [31].

Statistical analysis

All statistical analysis was done in R (v 3.4.3). Bland–Altman statistics were used to describe the bias and limits of agreement (LoA) of the accuracy and short-term repeatability for T_1, T_2, and ADC. Kendall's Tau (two-sided) was used to identify dependencies of the variation on the mean value, with a significance level of $\alpha = 0.05$. For parameters with significant dependencies, the relative percent ratio instead of the differences was used for the y-axis of the Bland–Altman plots [32,33]. For accuracy, the difference between the measured and the reference values was plotted as function of the reference value.

Additionally, the accuracy was calculated for each individual tube as the absolute percentage:

$$\textbf{Accuracy} = \frac{|\textbf{Measured} - \textbf{Reference}|}{\textbf{Reference}} \cdot 100\% \quad (1)$$

with the reference values specified by the phantom manufacturers. For short-term repeatability, Bland–Altman plots were produced by plotting the difference between the first and second intra-session measurements as a function of the mean of these two values.

Additionally, the repeatability for each individual phantom tube was calculated as follows:

$$\textbf{Repeatability} = \frac{|\textbf{Measurement } 2 - \textbf{Measurement } 1|}{\textbf{Mean}(\textbf{Measurement } 1, \textbf{Measurement } 2)} \cdot 100\% \quad (2)$$

Short-term repeatability was calculated using the repeated measurements of a single session. Long-term repeatability was calculated using Eq. (2) with repeat measurements on MR-linac A with five months in between.

Reproducibility (variation across systems) was quantified for T_1, T_2, and ADC mapping with the % coefficient of variation (%CV):

$$\%\textbf{CV} = \frac{\textbf{SD}(\textbf{system A}, \textbf{B}, \textbf{C}, \textbf{D})}{\textbf{Mean}(\textbf{system A}, \textbf{B}, \textbf{C}, \textbf{D})} \cdot 100\% \quad (3)$$

where the standard deviation (SD) and mean of the first measurements on each separate system were used.

Additional measures were calculated to enable comparison with previous studies. For T_2, the %CV was calculated for both short- and long-term repeatability, using the SD and mean of the repeat measurements. For the ADC maps, the %CV was calculated by using the SD and mean of the ROI in the tube in the isocenter of the systems. This was determined in the first acquisition of each measurement series. Additionally, diffusion images were analyzed according to the QIBA DWI profile [22].

For DCE, the stability was determined as the %CV, calculated with the ROI means and SDs over the 66–73 remaining scans from the five-minute acquisition.

For all acquisitions, only the slice at the center of the phantom was analyzed.

Results

An overview of the accuracy and short-term repeatability is provided in Fig. 2. The bias for the accuracy of T_1 (Fig. 2a) was found to be 11 ms with LoA of ±238 ms. For the T_1 short-term repeatability (Fig. 2b), the bias was −6 ms, and the LoA were ±63 ms. The accuracy, short-term repeatability, long-term repeatability, and reproducibility as calculated according to Eqs. (1)–(3) are presented in Table 1. Except for MR-linac B, the short-term repeatability was found to be lower than the accuracy.

The variation in the accuracy of T_2 (Fig. 2c) showed a dependence on the T_2 value ($\tau = .69, \boldsymbol{p} < .001$). This was also found for the T_2 repeatability ($\tau = .44, \boldsymbol{p} < .001$) (Fig. 2d). The bias and LoA, for T_2 were −11 ± 6 and 0 ± 2% for accuracy and short-term repeatability, respectively. The individual values for each system (Table 1) are comparable among systems for both accuracy and short-term repeatability, although the short-term repeatability is much lower.

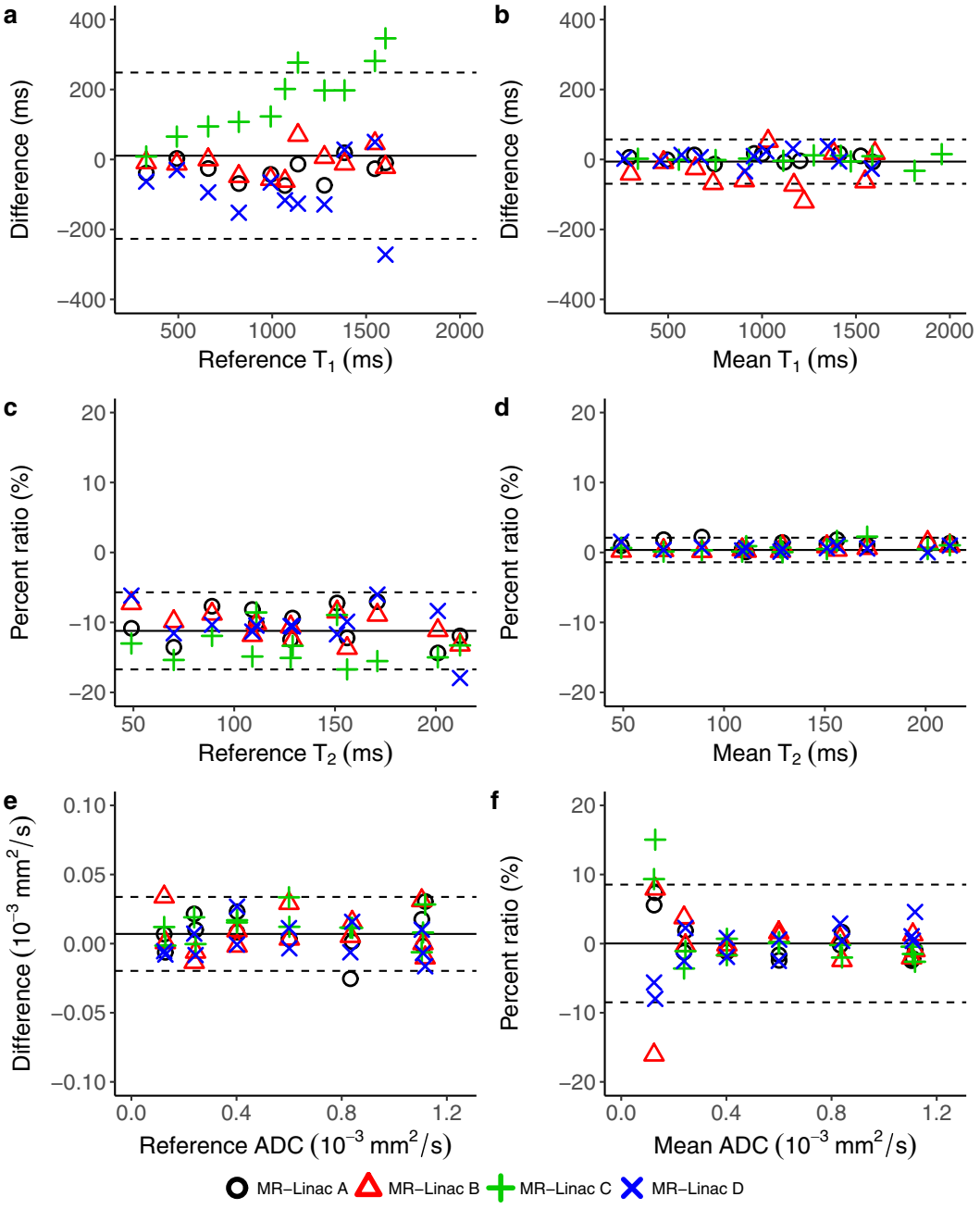

Fig. 2. Bland–Altman plots showing the accuracy and repeatability of T_1, T_2, and ADC mapping on the MR-linacs. The repeatability shown here (b,d,f) is the short-term repeatability, where two measurements were done consecutively. Each marker corresponds to a measurement in a single tube ROI. The solid line corresponds to the bias, and the dotted lines correspond to the LoA. Note that the *y*-axis of a, b, and e represent the qMRI units, while c, d, and f represent a percent ratio.

The variation in ADC repeatability (Fig. 2f) was found to depend on the measured ADC value ($\tau = .20$, $\boldsymbol{p} = .04$), so the ratio is shown. Bias and LoA for the accuracy (Fig. 2e) were $0.007 \cdot 10^{-3}$ mm^2/s, and $\pm 0.027 \cdot 10^{-3}$ mm^2/s. For the short-term repeatability of ADC (Fig. 2f), the bias was 0 and the LoA ± 9%. Individual values for the systems (Table 1) are similar for both accuracy and short-term repeatability, which in turn are comparable to each other.

The accuracy in the tube at the iso-center, which is also a measure described in the QIBA DWI profile, was found to be 0.2%, 0.0%, 0.7%, and 0.6% for MR-linac A, B, C, and D, respectively. The %CV based on the ROI mean and SD in the center tube were found to be 5% for MR-linac A, 9% for MR-linac B, 7% for MR-linac C, and 7% for MR-linac D. A more complete set of the QIBA DWI profile requirement measures can be found in Table S8, as well as results from the clinical protocol with larger voxels.

The DCE stability measurements produced a median %CV, which represents the variation in the signal intensity over the five-minute acquisition, of 0.6 (range: 0.2–2.0) % over the 10 tubes in MR-linac A. In MR-linac B the median %CV was 0.1 (range: 0.0–1.8) %, in MR-linac C 0.1 (range: 0.0–2.5) %, and in MR-linac D 0.6 (range: 0.2–2.9) %. Fig. 3 shows an increasing deviation from the reference value with an increase in concentration. For concentrations of 0.5 mM, the median accuracy was 23.5 (range: 14.8–35.5) %. For higher concentrations, this increased to a median accuracy of 62.0 (range: 47.6–71.2) % for 9.8 mM. Long-term repeatability was found to be 17.9 (median, range: 1.0–37.9) %, and repro-

Table 1
Accuracy and repeatability of T_1, T_2, and ADC mapping. For each system and qMRI parameter, the median (range) of the measured phantom tubes is given.

	T_1 (%)	T_2 (%)		ADC (%)
Accuracy				
MR-linac A	4.0 (0.6–11.8)	10.5 (7.0–14.4)		2.7 (0.2–9.0)
MR-linac B	2.7 (0.1–6.1)	10.4 (7.3–13.7)		1.9 (0.0–27.1)
MR-linac C	14.3 (2.6–24.4)	14.1 (8.6–16.7)		2.0 (0.1–9.6)
MR-linac D	10.9 (2.0–19.2)	10.5 (6.0–18.0)		1.9 (0.2–6.6)
Short-term repeatability			%CV	
MR-linac A	1.2 (0.2–2.1)	1.2 (0.1–2.2)	0.8 (0.1–1.6)	1.7 (0.2–7.1)
MR-linac B	5.1 (1.1–13.8)	0.4 (0.2–1.5)	0.3 (0.1–1.0)	1.3 (0.2–17.5)
MR-linac C	0.6 (0.1–1.8)	0.6 (0.1–2.2)	0.4 (0.0–1.6)	1.5 (0.1–14.0)
MR-linac D	1.7 (0.4–3.7)	0.6 (0.0–1.5)	0.4 (0.0–1.1)	2.2 (0.4–8.3)
Long-term repeatability				
MR-Linac A	1.8 (0.4–5.6)	1.4 (0.8–2.9)	1.0 (0.6–2.1)	1.7 (0.0–6.4)
Reproducibility (%CV)				
All systems	11.2 (6.6–15.8)	2.9 (0.9–4.7)		2.2 (0.6–12.0)

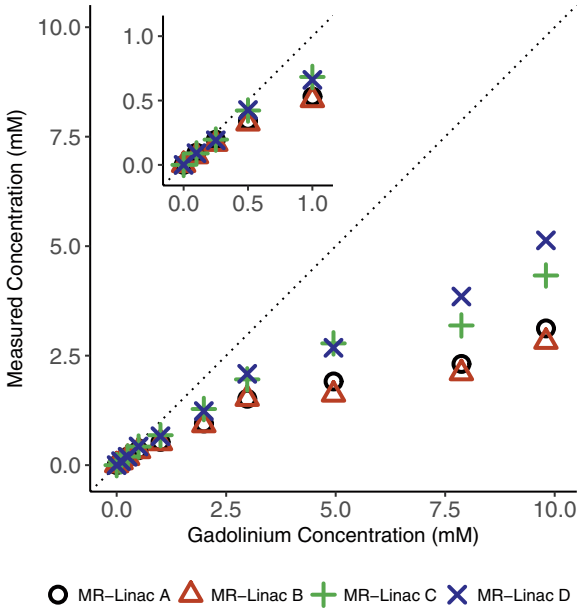

Fig. 3. Measured gadolinium concentration plotted against the reference value. The dotted line represents the identity line ($y = x$).

ducibility (%CV) was 16.7 (median, range: 8.0–28.3) %, which is high compared to the other modalities (Table 1).

Fig. 4 shows the quantitative maps from one prostate cancer patient (62 years, initial PSA 37 µg/L, Gleason score 4+5), acquired before the start of radiation treatment. The tumor is clearly visible on the T_2, ADC, and K^{trans} maps (Fig. 4b–d), and the images indicate good quality with minimal distortions and no obvious artefacts.

Discussion

MR-linac systems may enable daily qMRI acquisitions for treatment response monitoring, prediction, and biomarker discovery during radiotherapy. In this study, we determined the accuracy, repeatability, and reproducibility of quantitative T_1 mapping, T_2 mapping, ADC mapping, and DCE-MRI on the Unity MR-linac. Additionally, we are the first to demonstrate feasibility of these quantitative acquisitions on the Unity MR-linac in a patient with prostate cancer.

The Bland–Altman plots provide an overview of the four MR-linacs, which means that the reported biases and LoAs summarize the group, and do not represent individual systems. For T_1 and T_2 they however show clearly that repeated measurements on a single system (Fig. 2b,d) show less variation than measurements between systems (Fig. 2a,c). The former is interesting for single-center studies, where one system is used to assess variations in individual patients over time, and the latter is interesting for multi-center studies.

The accuracy of the T_1 VFA series (Fig. 2a) shows great variation between systems, especially for higher T_1 values. This is likely the result of using the VFA method, which is known to depend on several parameters, e.g. spoiling, and is generally known to overestimate T_1 values [28,29]. The variation between systems underwrites that careful validation and optimization is needed if the VFA were to be used in a multi-center study.

For T_2 accuracy, a clear negative bias was found over the entire range of T_2 values, which was also found previously in studies using a multi-echo approach [34–36]. As this bias is over the entire range of T_2 values, this should be of little influence to detect differences in tissues. The LoA indicate a variation of 12%, which corresponds to 6–25 ms over the range of assessed T_2 values. Therefore, differences between for instance a prostate tumor (80 ms) and healthy prostate tissue (150 ms) [37,38] should be very well detectable. The zero bias and narrow LoA of the repeatability indicate that small changes in T_2 due to radiation treatment should be detectable. Short- and long-term repeatability %CV are comparable to the results of diagnostic systems [38].

For ADC mapping, the QIBA DWI profile presents threshold values for measurements in the iso-center that represent requirements for systems to meet the profile claims about confidence intervals for ADC measurements in patients [22]. For the accuracy, all systems passed the requirement of $\leq$3.6%, as the highest value found was 0.7%, indicating that ADC measurements in a ROI are accurate. This center tube accuracy also compares to previously found values in diagnostic systems [39,40]. On the other hand, none of the systems met the requirement for the %CV, as this was above the recommended 2% for all systems. This was also found previously on a 1.5T system [41]. One reason for the increased ROI %CV values is the reduced signal to noise ratio of the available 8-channel body array [18] compared to a head coil as recommended by the QIBA. The %CV is determined using the SD of the ROI in the central tube and is therefore closely related to the SNR. Indeed, evaluation of the clinical acquisition sequence with larger voxels and lower maximum b-value showed that the % CV requirements were met (Table S8). For the other tubes, the accuracy deteriorated up to 27.1% in MR-linac B for the vial with the lowest ADC ($0.125 \cdot 10^{-3}$mm^2/s). The repeatability also worsened at these low ADC values. These deviations can partly be explained by lower SNR due to the receiver coil, but possibly also by gradient nonlinearities which influence ADC values measured

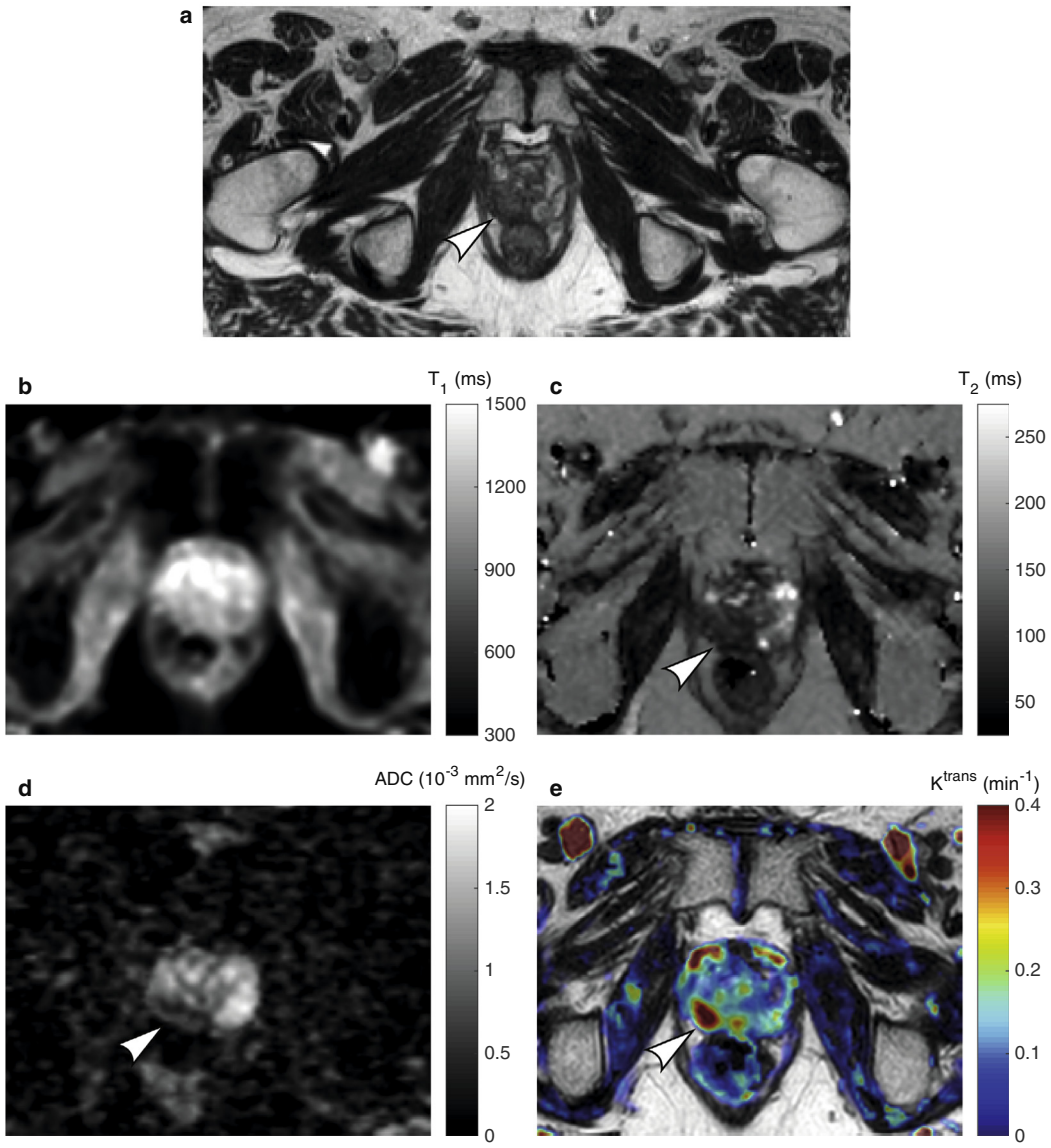

Fig. 4. Quantitative maps of a patient with prostate cancer. (a) T_2-weighted anatomical image, (b) T_1 map, (c) T_2 map, (d) ADC map, (e) K^{trans} map. The white arrows indicate the location of the tumor, visible on the T_2 map, ADC map, and K^{trans} map.

away from the iso-center [39]. This influence, as well as system based geometric distortions of the Unity MR-linac should be quantified, especially if the goal for imaging is treatment planning or dose painting [42]. Short- and long-term repeatability of ADC mapping in the iso-center are comparable to previously published results [39].

For DCE, good stability was found, with a %CV in signal intensity below 2% over all tubes over the course of five minutes. The gadolinium concentrations were determined inaccurately, where concentrations of 0.5 mM could be estimated within a range of ~30%. For higher gadolinium concentrations, the systems all show a severe underestimation. This should not be a problem for low-perfused organs like the prostate but might result in an underestimation for well-perfused tissues. Although the errors in repeatability and reproducibility are relatively high, the patient K^{trans} image shows the added value on the single patient level, as the tumor is clearly visible (Fig. 4e).

In conclusion, we assessed the performance of the Unity MR-linac for a range of quantitative MR sequences and showed the feasibility of qMRI in a single patient. The accuracy and repeatability for T1 and T2 are similar to literature values from diagnostic systems. ADC mapping is also accurate although larger voxels might be advisable to increase the SNR. DCE acquisitions are decreasingly accurate for increasing contrast agent concentration but are stable and valuable for individual patients. This indicates that the Unity MR-linac performs similar to diagnostic MRI systems and can be used for treatment response monitoring and biomarker discovery studies.

Declaration of interest

The Netherlands Cancer Institute, University Medical Center Utrecht, Medical College of Wisconsin, The Institute of Cancer Research, and The University of Texas MD Anderson Cancer Center are members of the Elekta MR-linac Consortium, which aims to coordinate international collaborative research relating to the Elekta Unity (MR-linac). Elekta and Philips are commercial partners within the Consortium. Elekta financially supports consortium member institutions with research funding and travel costs for Consortium meetings.

The Institute of Cancer Research (ICR) is supported by Cancer Research UK under programme C33589/A19727 and C7224/

A23275. This paper represents independent research partly funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at the Royal Marsden NHS Foundation Trust and the ICR. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Dr. van der Heide is supported by a grant from ITEA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2019.01.011.

References

[1] Zahra MA, Hollingsworth KG, Sala E, Lomas DJ, Tan LT. Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. Lancet Oncol 2007;8:63–74. https://doi.org/10.1016/S1470-2045(06)71012-9.

[2] Iima M, Le Bihan D. Clinical intravoxel incoherent motion and diffusion MR imaging: past, present, and future. Radiology 2016;278:13–32. https://doi.org/10.1148/radiol.2015150244.

[3] Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS prostate imaging-reporting and data system: 2015, Version 2. Eur Urol 2016;69:16–40. https://doi.org/10.1016/j.eururo.2015.08.052.

[4] Morris EA, Comstock CE, Lee CH. ACR BI-RADS® magnetic resonance imaging. ACR BI-RADS® Atlas, Breast Imaging Report. Data Syst.. Reston, VA: American College of Radiology; 2013.

[5] O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14:169–86. https://doi.org/10.1038/nrclinonc.2016.162.

[6] Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. Stat Methods Med Res 2015;24:27–67. https://doi.org/10.1177/0962280214537344.

[7] Keenan KE, Ainslie M, Barker AJ, Boss MA, Cecil KM, Charles C, et al. Quantitative magnetic resonance imaging phantoms: a review and the need for a system phantom. Magn Reson Med 2017. https://doi.org/10.1002/mrm.26982.

[8] Decker G, Mürtz P, Gieseke J, Träber F, Block W, Sprinkart AM, et al. Intensity-modulated radiotherapy of the prostate: dynamic ADC monitoring by DWI at 3.0 T. Radiother Oncol 2014;113:115–20. https://doi.org/10.1016/j.radonc.2014.07.016.

[9] Ho JC, Allen PK, Bhosale PR, Rauch GM, Fuller CD, Mohamed ASR, et al. Diffusion-weighted magnetic resonance imaging as a predictor of outcome in cervical cancer after chemoradiation. Int J Radiat Oncol Biol Phys 2017;97:546–53. https://doi.org/10.1016/j.ijrobp.2016.11.015.

[10] Das S, Chandramohan A, Rami Reddy JK, Mukhopadhyay S, Kumar RM, Isiah R, et al. Role of conventional and diffusion weighted MRI in predicting treatment response after low dose radiation and chemotherapy in locally advanced carcinoma cervix. Radiother Oncol 2015;117:288–93. https://doi.org/10.1016/j.radonc.2015.10.006.

[11] Wang L, Liu L, Han C, Liu S, Tian H, Li Z, et al. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. Radiother Oncol 2016;121:246–51. https://doi.org/10.1016/j.radonc.2016.10.021.

[12] Mahmood F, Johannesen HH, Geertsen P, Hansen RH. Repeated diffusion MRI reveals earliest time point for stratification of radiotherapy response in brain metastases. Phys Med Biol 2017;62:2990–3002. https://doi.org/10.1088/1361-6560/aa5249.

[13] Datta A, Aznar MC, Dubec M, Parker GJM, O'Connor JPB. Delivering functional imaging on the MRI-linac: current challenges and potential solutions. Clin Oncol 2018;30:702–10. https://doi.org/10.1016/j.clon.2018.08.005.

[14] Raaymakers BW, Lagendijk JJW, Overweg J, Kok JGM, Raaijmakers AJE, Kerkhof EM, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. Phys Med Biol 2009;54:N229–37. https://doi.org/10.1088/0031-9155/54/12/N01.

[15] Mutic S, Dempsey JF. The ViewRay system: magnetic resonance-guided and controlled radiotherapy. Semin Radiat Oncol 2014;24:196–9. https://doi.org/10.1016/j.semradonc.2014.02.008.

[16] Fallone BG. The rotating biplanar linac-magnetic resonance imaging system. Semin Radiat Oncol 2014;24:200–2. https://doi.org/10.1016/j.semradonc.2014.02.011.

[17] Keall PJ, Barton M, Crozier S. The Australian magnetic resonance imaging-linac program. Semin Radiat Oncol 2014;24:203–6. https://doi.org/10.1016/j.semradonc.2014.02.015.

[18] Hoogcarspel SJ, Zijlema SE, Tijssen RHN, Kerkmeijer LGW, Jürgenliemk-Schulz IM, Lagendijk JJW, et al. Characterization of the first RF coil dedicated to 1.5 T MR guided radiotherapy. Phys Med Biol 2018;63. https://doi.org/10.1088/1361-6560/aaa303.

[19] Wojcieszynski AP, Rosenberg SA, Brower JV, Hullett CR, Geurts MW, Labby ZE, et al. Gadoxetate for direct tumor therapy and tracking with real-time MRI-guided stereotactic body radiation therapy of the liver. Radiother Oncol 2016;118:416–8. https://doi.org/10.1016/j.radonc.2015.10.024.

[20] Yang Y, Cao M, Sheng K, Gao Y, Chen A, Kamrava M, et al. Longitudinal diffusion MRI for treatment response assessment: preliminary experience using an MRI-guided tri-cobalt 60 radiotherapy system. Med Phys 2016;43:1369–73. https://doi.org/10.1118/1.4942381.

[21] Gao Y, Han F, Zhou Z, Cao M, Kaprealian T, Kamrava M, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: sequence verification and preliminary clinical experience. Med Phys 2017;44:5357–66. https://doi.org/10.1002/mp.12465.

[22] Perfusion, Diffusion and Flow-MRI Biomarker Committee. Diffusion-Weighted Magnetic Resonance Imaging (DWI), Quantitative Imaging Biomarkers Alliance. Initial Draft. QIBA 2017.

[23] DCE MRI Technical Committee. DCE MRI quantification profile, quantitative imaging biomarkers alliance. Version 1.0. Reviewed draft. QIBA; 2012.

[24] Fram EK, Herfkens RJ, Johnson GA, Glover GH, Karis JP, Shimakawa A, et al. Rapid calculation of T1 using variable flip angle gradient refocused imaging. Magn Reson Imaging 1987;5:201–8. https://doi.org/10.1016/0730-725X(87)90021-X.

[25] Deoni SCL, Rutt BK, Peters TM. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. Magn Reson Med 2003;49:515–26. https://doi.org/10.1002/mrm.10407.

[26] Milford D, Rosbach N, Bendszus M, Heiland S. Mono-exponential fitting in T2-relaxometry: relevance of offset and first echo. PLoS ONE 2015;10:1–13. https://doi.org/10.1371/journal.pone.0145255.

[27] Schabel MC, Parker DL. Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences. Phys Med Biol 2008;53:2345–73. https://doi.org/10.1088/0031-9155/53/9/010.

[28] Bane O, Hectors SJ, Wagner M, Arlinghaus LL, Aryal MP, Cao Y, et al. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. Magn Reson Med 2017;00:1–12. https://doi.org/10.1002/mrm.26903.

[29] Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the accuracy of T1 mapping: searching for common ground. Magn Reson Med 2015;73:514–22. https://doi.org/10.1002/mrm.25135.

[30] Murase K. Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging. Magn Reson Med 2004;51:858–62. https://doi.org/10.1002/mrm.20022.

[31] Parker GJM, Roberts C, Macdonald A, Buonaccorsi GA, Cheung S, Buckley DL, 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. https://doi.org/10.1002/mrm.21066.

[32] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60. https://doi.org/10.1177/096228029900800204.

[33] Bland JM, Altman DG. Statistics notes: measurement error proportional to the mean. BMJ 1996;313:106. https://doi.org/10.1136/bmj.313.7049.106.

[34] Andersen C, Tågehøj Jensen F. Precision, accuracy, and image plane uniformity in NMR relaxation time imaging on a 1.5 T whole-body MR imaging system. Magn Reson Imaging 1994;12:775–84. https://doi.org/10.1016/0730-725X(94)92202-0.

[35] Graves MJ, Emmens D, Lejay H, Hariharan H, Polzin J, Lomas DJ. T2 and T2* quantification using optimal B1 image reconstruction for multicoil arrays. J Magn Reson Imaging 2008;28:278–81. https://doi.org/10.1002/jmri.21420.

[36] Liney GP, Knowles AJ, Manton DJ, Turnbull LW, Blackband SJ, Horsman A. Comparison of conventional single echo and multi-echo sequences with a Fast Spin-Echo sequence for quantitative T2 mapping: application to the prostate. J Magn Reson Imaging 1996;6:603–7. https://doi.org/10.1002/jmri.1880060408.

[37] de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DCMR. Imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. Radiology 2004;230:652–9. https://doi.org/10.1148/radiol.2303021331.

[38] van Houdt PJ, Agarwal HK, van Buuren LD, Heijmink SWTPJ, Haack S, van der Poel HG, et al. Performance of a fast and high-resolution multi-echo spin-echo sequence for prostate T2 mapping across multiple systems. Magn Reson Med 2017;79:1586–94. https://doi.org/10.1002/mrm.26816.

[39] Malyarenko D, Galbán CJ, Londy FJ, Meyer CR, Johnson TD, Rehemtulla A, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. J Magn Reson Imaging 2013;37:1238–46. https://doi.org/10.1002/jmri.23825.

[40] Grech-Sollars M, Hales PW, Miyazaki K, Raschke F, Rodriguez D, Wilson M, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. NMR Biomed 2015;28:468–85. https://doi.org/10.1002/nbm.3269.

[41] Moreau B, Iannessi A, Hoog C, Beaumont H. How reliable are ADC measurements? A phantom and clinical study of cervical lymph nodes. Eur Radiol 2018;28:3362–71. https://doi.org/10.1007/s00330-017-5265-2.

[42] Schmidt MA, Payne GS. Radiotherapy planning using MRI. Phys Med Biol 2015;60:R323–61. https://doi.org/10.1088/0031-9155/60/22/R323.