Validation of T2 mapping for treatment response monitoring in longitudinal multi-center clinical trials

Petra J van Houdt[1], Harsh K Agarwal[2,3], Laurens B van Buuren[1], Marko Ivancevic[4], Søren Haack[5], Jesper Folsted Kallehauge[6], Peter L Choyke[3], and Uulke A van der Heide[1]

[1]Radiation Oncology, the Netherlands Cancer Institute, Amsterdam, Netherlands, [2]Philips Research NA, Briarcliff Manor, MD, United States, [3]National Cancer Institute, National Institutes of Health, Bethesda, NY, United States, [4]Philips Healthcare, Best, Netherlands, [5]Clinical Engineering, Aarhus University Hospital, Aarhus, Denmark, [6]Medical Physics, Aarhus University Hospital, Aarhus, Denmark

Purpose T2 is proposed as a potential biomarker for response monitoring of patients with prostate cancer treated with radiotherapy [1]. Recently an accelerated multi-echo spin-echo (ME-SE) T2 mapping technique (k-t T2) has been validated for whole-prostate imaging in a clinically reasonable time without the loss of spatial resolution or dynamic range [2,3]. A key requirement in multi-center longitudinal studies is good reproducibility over multiple visits. Therefore, the purpose of this study was to establish the intra- and inter-scanner reproducibility of T2 maps using k-t T2 technique.

Methods Validation of the k-t T2 sequence was performed on three clinical MR platforms, one 1.5 T system (A) and two 3T systems (B and C) at two different institutes.

Protocol We optimized the k-t T2 sequences as proposed in [3] for each platform to account for differences in scanner hardware and receiver coils. Spatial resolution and FOV were adjusted to obtain whole prostate maps at comparable SNR within clinically feasible scan times (< 6min). Preliminary experiments at platform B revealed a substantial difference (6.4 ± 1.1%) between T2 values estimated from un-accelerated ME-SE scans with perfect and imperfect (i.e. truncated) refocusing pulses. The use of perfect refocusing pulses restricts minimal TE, which varies between platforms. Therefore, we chose the settings for FOV, voxel size, TR, TE, no. of echoes, and maximum allowed B1 transmit field in such a way that perfect refocusing pulses are used while keeping the same TE dynamic range and good SNR images (Table 1). Other settings were identical: fold-over suppression with three times FOV; partial-Fourier factor = 0.625, SENSE factor = 2, k-t T2 reduction factor = 4.

*Table 1: Platform specific settings (*with endocoil)*

	Platform A	Platform B	Platform C
No. echoes	16	12	12
TE1/ ΔTE (ms)	22/11	32/16	32/16
TR (ms)	2693	3200/ 2459*	3196/2273*
Max B1 field (μT)	maximum	9.5	9.5
Voxel size (mm^3)	1.5 x 1.5 x 3	1.2 x1.2 x 3/ 0.8 x 0.8 x 3*	1.1 x 1.1 x 3/ 0.6 x 0.6 x 3*
FOV (mm^3)	170 x 170 x 60	170 x 170 x 60	210 x 210 x 60/ 170 x 170 x 60*

Phantom measurements A Eurospin T05 phantom, consisting of gel samples with varying T1 and T2 values was used (Diagnostic Sonar, Livingston, Scotland). We chose a set of 12 gels with T2 values in the expected range of healthy and tumorous prostate tissue. T2 maps were generated by fitting a mono-exponential decay function for each voxel with a nonlinear least-squares method. Average T2 values were determined for a ROI centered in each tube (diameter 1.2 cm). T2 values were all normalized to the same temperature (296 K) by measuring the temperature before and after each scan.

First, we tested the effect of acceleration on T2 accuracy by repeating the k-t T2 sequence ten times within one scan session and comparing the T2 values ($T2_{k\text{-}t2}$) to reference T2 values ($T2_{ref}$). The latter were obtained from the same ME-SE sequence with perfect pulses and similar dynamic range, but without any acceleration. The relative difference between the T2 values (ΔT2) from the k-t T2 sequence and the reference sequence is a measure for the accuracy of the k-t T2 sequence. Second, inter-scanner variability was assessed by comparing the T2 values from platform B to A. The use of a phantom from a different batch in platform C prohibited a comparison with platforms A and B. Third, long-term reproducibility was assessed by repeating k-t T2 scans ten times once every month over a period of six months on platforms A and B. We compared the T2 values of each measurement to the average T2 values ($T2_{avg}$) across all 60 measurements. This would identify differences due to day-to-day variations, minor software upgrades, and repositioning of the phantom, which are changes that we also expect to happen during clinical trials.

Volunteers Repeated T2 maps were acquired on two different days for five volunteers using k-t T2 sequence. The scans were performed on platform B with a 6-channel cardiac coil. T2 values were estimated with maximum likelihood estimation to account for Rician noise characteristics [4].

Results Fig. 1 shows the accuracy of the k-t T2 sequence compared to the reference ME-SE sequence for the three platforms. A systematic error of -1.1 ± 0.9% was observed for platform A, 0.0 ± 1.1% for platform B, and 1.4 ± 0.8% for platform C. Inter-scanner variability between platform A and B was 0.0 ± 1.7%. Long-term reproducibility results are shown in Fig. 2; T2 values varied between -0.6 ± 1.1% for platform A and 0.5 ± 1.2% for platform B for the six month period. Table 2 shows the average T2 values for the whole prostate for all volunteers. Differences in average prostate T2 values between the two scans are -1.5 ± 9.5% across all volunteers. Fig. 3 shows an example of prostate T2 maps of two different days.

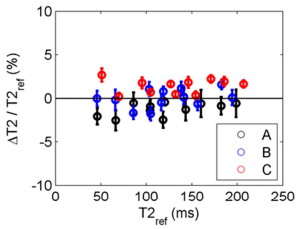

*Fig. 1: Accuracy of k-t T2. Each marker represents a gel sample. Error bars = 2*std for repeated measurements (n=10).*

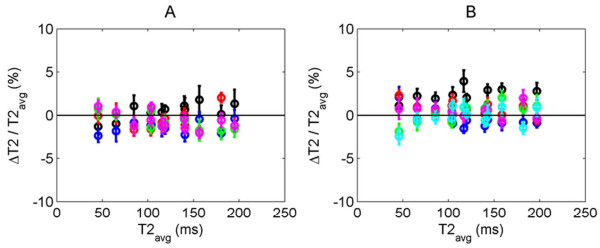

*Fig. 2: Long-term reproducibility for platform A and B. Each marker represents a gel sample on a given day. Different colors represent measurements on different days (n=6). Error bars = 2*std for repeated measurements on a given day (n=10).*

Table 2: Average T2 values of whole prostate

Volunteer	Day 1 (ms)	Day 2 (ms)
1	102.4 ± 38.6	95.9 ± 36.6
2	136.5 ± 57.6	149.6 ± 67.5
3	119.3 ± 39.5	111.0 ± 31.8
4	103.3 ± 37.1	118.8 ± 51.6
5	112.5 ± 37.8	110.8 ± 38.4

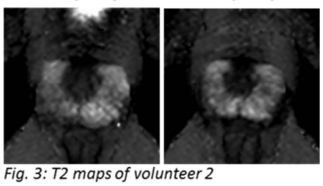

Fig. 3: T2 maps of volunteer 2

Discussion and conclusion The differences in T2 values within and between platforms are smaller than 5%. The acceleration techniques do not affect the accuracy in T2 values. The differences in T2 values for volunteers are on average larger than the reproducibility from phantom data most likely due to lower signal-to-noise, different coil loading, and motion. These data suggest that T2 mapping can be used to distinguish tumorous (~80 ms) and healthy prostate tissue (~115 ms). More test-retest volunteer or patient data including a voxel-based repeatability analysis are required to determine the smallest detectable difference in T2 values that can be attributed to a treatment effect. In conclusion, this work demonstrates that T2 mapping using k-t T2 technique is suitable for use in multi-center, longitudinal trials in prostate cancer patients. For such multi-center studies, a dedicated quality assurance program is required to test the T2 accuracy and reproducibility of each participating scanner [5]. The results of this study are also applicable for other types of cancer for which the expected T2 values are in the same range.

Acknowledgements This study is part of the DR THERAPAT project (FP7-ICT-2011-9, Project No. 600852).

References [1] Foltz et al. JMRI 2013; 37:909-916 [2] Liu et al. Magn Reson Med 2011; 65:1400-1406 [3] Agarwal et al. ISMRM 2013; 1373 [4] Bos et al. ISMRM 2009; 4526 [5] Tofts et al. BJR 2011; 84:S213-26.