Improvements in Between-Vendor MRI Harmonization of Renal T_2 Mapping using Stimulated Echo Compensation

Hao Li, PhD,[1,2] Alexander J. Daniel, PhD,[3] Charlotte E. Buchanan, PhD,[3] Fábio Nery, PhD,[4] David M. Morris, PhD,[5] Shaohang Li, BS,[1] Yuan Huang, PhD,[2,6] João A. Sousa, MS,[7] Steven Sourbron, PhD,[7] Iosif A. Mendichovszky, MD,[2,8] David L. Thomas, PhD,[9,10] Andrew N. Priest, DPhil,[2,8] and Susan T. Francis, PhD[3,11*]

Background: T_2 mapping is valuable to evaluate pathophysiology in kidney disease. However, variations in T_2 relaxation time measurements across MR scanners and vendors may occur requiring additional correction.
Purpose: To harmonize renal T_2 measurements between MR vendor platforms, and use an extended-phase-graph-based fitting method ("StimFit") to correct stimulated echoes and reduce between-vendor variations.
Study Type: Prospective.
Subjects: 8 healthy "travelling" volunteers (37.5% female, 32 ± 6 years) imaged on four MRI systems across three vendors at four sites, 10 healthy volunteers (50% female, 32 ± 8 years) scanned multiple times on a given MR scanner for repeatability evaluation. ISMRM/NIST system phantom scanned for evaluation of T_2 accuracy.
Field Strength/Sequence: 3T, multiecho spin-echo sequence.
Assessment: T_2 images fit using conventional monoexponential fitting and "StimFit." Mean absolute percentage error (MAPE) of phantom measurements with reference T_2 values. Average cortex and medulla T_2 values compared between MR vendors, with masks obtained from T_2-weighted images and T_1 maps. Full-width-at-half-maximum (FWHM) T_2 distributions to evaluate local homogeneity of measurements.
Statistical Tests: Coefficient of variation (CV), linear mixed-effects model, analysis of variance, student's *t*-tests, Bland–Altman plots, *P*-value <0.05 considered statistically significant.
Results: In the ISMRM/NIST phantom, "StimFit" reduced the MAPE from 4.9%, 9.1%, 24.4%, and 18.1% for the four sites (three vendors) to 3.3%, 3.0%, 6.6%, and 4.1%, respectively. In vivo, there was a significant difference in kidney T_2 measurements between vendors using a monoexponential fit, but not with "StimFit" ($P = 0.86$ and 0.92, cortex and medulla, respectively). The intervendor CVs of T_2 measures were reduced from 8.0% to 2.6% (cortex) and 7.1% to 2.8% (medulla) with StimFit, resulting in no significant differences for the CVs of intravendor repeat acquisitions ($P = 0.13$ and 0.05). "StimFit" significantly reduced the FWHM of T_2 distributions in the cortex and whole kidney.
Data Conclusion: Stimulated-echo correction reduces renal T_2 variation across MR vendor platforms.
Level of Evidence: 2
Technical Efficacy: Stage 1

J. MAGN. RESON. IMAGING 2024.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.29282

Received Nov 9, 2023, Accepted for publication Jan 23, 2024.

*Address reprint requests to: S.T.F., Room MR13, Sir Peter Mansfield Imaging Centre, University Park Nottingham, Nottinghamshire, NG7 2RD, UK.
E-mail: Susan.Francis@nottingham.ac.uk
The last two authors are Joint senior authors.

From the [1]The Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai, China; [2]Department of Radiology, University of Cambridge, Cambridge, UK; [3]Sir Peter Mansfield Imaging Centre, University of Nottingham, Nottingham, UK; [4]Developmental Imaging and Biophysics Section, UCL Great Ormond Street Institute of Child Health, London, UK; [5]Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK; [6]EPSRC Cambridge Mathematics of Information in Healthcare Hub, University of Cambridge, Cambridge, UK; [7]Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK; [8]Department of Radiology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK; [9]Neuroradiological Academic Unit, UCL Queen Square Institute of Neurology, University College London, London, UK; [10]Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, London, UK; and [11]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and School of Medicine, Nottingham, UK

MRI T_2 mapping is sensitive to edematous changes and ischemia.[1] In the kidney, it has shown potential in the evaluation of autosomal dominant polycystic kidney disease,[2] renal cell carcinomas,[3] ischemia–reperfusion injury,[4,5] and renal transplants.[6,7] Mapping of absolute T_2 values can potentially enable a more objective study of disease-related changes over time than T_2-weighted MRI. Although T_2 is an inherent tissue property, quantitative assessment of tissue T_2 relaxation time is dependent on various factors including pulse-sequence type, radiofrequency (RF) pulse profile, acquisition parameters, MRI hardware capabilities, and subject-specific influences of coil loading and transmit/receive gain settings.[8] The accuracy and reproducibility of T_2 measurements should be investigated, particularly when combining data across MR vendors and platforms.[9]

Quantitative T_2 maps can be acquired using various pulse sequences, including multiecho spin echo (MESE), gradient and spin echo,[10] T_2-prepared single-shot balanced steady-state free precession,[11] and driven equilibrium single-pulse observation of T_2 sequences.[12] The MESE pulse sequence is widely used due to its commercial availability across all MR vendors.[13] However, challenges arise from B_1 field inhomogeneities, imperfect slice selection pulse profiles, and transmit calibration errors, causing deviations from the nominal 180° flip angle refocusing pulses.[8] The resulting stimulated and indirect echoes cause T_2 values to be overestimated,[14] particularly for body imaging at 3T. This bias can vary between scanners with different hardware, RF pulse shapes, and protocol imlementations,[15] which is problematic for any multicenter clinical trials.

The effects of indirect echoes can be corrected by postprocessing methods, such as discarding particular echoes from the fit,[16] model-based methods using the extended phase graph (EPG) algorithm,[8,14,17] and dictionary-based methods.[18–20] These methods have been reported to be effective in phantom and in vivo measurements, but only for studies using MR scanners of single vendors or with investigations at a single imaging site.[8,14,17–20] The effects of indirect echoes on multivendor and multicenter performance, and whether such biases can be corrected, remains unexplored.

The United Kingdom Renal Imaging Network-MRI acquisition and processing standardization (UKRIN-MAPS) project[21,22] was set up to develop harmonized renal MRI protocols across MR vendors, which are in-line with the recent consensus guidelines regarding patient preparation, hardware, acquisition parameters (for T_2 mapping: >5 echo times, maximum echo time >120 msec at 3T) and data analysis.[13,23] A preliminary investigation found a large cross-vendor variation in renal T_2 when using a monoexponential fit, despite using a standardized MESE sequence across MR vendors with harmonized parameters.[15]

This study aims to evaluate the consistency of renal T_2 measurements obtained across 3T MR platforms from different vendors (GE, Philips, and Siemens) using an EPG-based fitting method.

Materials and Methods

This study was a cross-site study with MRI data collected at four imaging sites (Sir Peter Mansfield Imaging Centre, University of Nottingham; Department of Radiology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust; Developmental Imaging and Biophysics Section, Great Ormond Street Institute of Child Health, University College London; Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK) with participants scanned under healthy volunteer ethics approval from the local research ethics boards. All participants provided written informed consent.

MRI Data Acquisition

Experiments were performed at 3T on four MRI systems from three different vendors (Discovery MR750, General Electric [GE] Healthcare, Waukesha, WI, USA; Ingenia, Philips Healthcare, Best, Netherlands; two MRI systems tested-Prisma and Skyra-Fit, Siemens Healthcare, Erlangen, Germany) at four imaging sites. Scanners were equipped with a dual-channel transmit system, except for the GE Discovery MR 750 that used a single-channel system.

A respiratory-triggered MESE sequence was harmonized across vendors as part of the UKRIN-MAPS renal MRI protocol.[21] Key parameters included a minimum repetition time (TR) = 3 sec, echo time (TE) = 12.9–129.0 msec in 12.9 msec steps, nominal refocusing flip angle = 180°, field of view (FOV) = 38.4 cm, acquisition matrix = 128 × 128, five slices with thickness/gap = 4.5/1.0 mm, parallel imaging factor = 3, and acquisition time = 43 breaths. The approximate acquisition time for collection of the T_2 mapping was 3 minutes dependent on breathing rate. The GE product MESE sequence was customized to enable controllable echo spacing. Detailed parameters for UKRIN-MAPS and National Institute of Standards and Technology (NIST) reference protocols are shown in Table 1.

A B_1^+ map was also collected on each scanner in matched native space to the MESE acquisition using the vendor-specific B_1 mapping scheme (Philips: Dual Refocusing Echo Acquisition Mode (DREAM[24]); Siemens: TurboFLASH B_1 mapping[25]; GE: Bloch-Siegert method[26]) to evaluate T_2 variations caused by B_1 inhomogeneity and to compare with B_1 maps estimated by the EPG "StimFit" model.[8,27] In addition, a harmonized T2-weighted single-shot fast-spin-echo sequence[15] (TE = 60 msec, TR = 900–1300 msec, SENSE/ASSET/GRAPPA = 3/3/2, refocus angle 120°, bandwidth, 792 Hz, voxel size = 1.5 × 1.5 × 5 mm^3 with slice gap 0.5 mm, 17 coronal slices, in a single 15–17 sec breath-hold) and a modified look-locker imaging (MOLLI) T1 mapping sequence[28] with a 5(3)3 acquisition scheme (TE 1.12–1.55 msec, flip angle 20°, voxel size = 1.5 × 1.5 × 5 mm^3 with slice gap 0.5 mm, 5 slices, 1 BH per slice) were acquired in matched native space. Full MESE, B_1 mapping, T_2-weighted and MOLLI sequence parameters can be downloaded from https://www.nottingham.ac.uk/research/groups/spmic/research/uk-renal-imaging-network/ukrin-maps.aspx.

Phantom Experiments

The International Society for Magnetic Resonance in Medicine/National Institute of Standards and Technology (ISMRM/NIST) system phantom[29] was used to evaluate the accuracy of T_2 measurements against the T_2-array reference values provided by the manufacturer. The phantom was scanned three times on each scanner

TABLE 1. Key Parameters of the NIST and UKRIN-MAPS T_2 Mapping Protocols

Protocol	NIST	UKRIN-MAPS
TE	GE & Siemens: 10 msec × 32 echoes[a] Philips: 11 msec × 16 echoes[b]	12.9 msec × 10 echoes
TR	5000 msec	Minimum: 3000 msec (phantom: 3750 msec)
Refocusing pulse flip angle	180°	180°
Resolution (mm)	1 × 1.3	3 × 3
Bandwidth	227 Hz/pixel	244 Hz/pixel
FOV	250 mm	384 mm
Slice number	1	5
Thickness/gap	6 mm	GE & Philips: 4.5/1 mm Siemens: 5/0.5 mm
Acceleration	None	Parallel imaging: ×3
Acquisition time	16 minutes 10 sec	43 breaths (phantom: 2 minutes 49 sec)

TE = echo time; TR = repetition time; NIST = National Institute of Standards and Technology; UKRIN-MAPS = UK renal imaging network-MRI acquisition and processing standardization.
[a]GE NIST reference T_2-mapping protocol was modified to a MESE to match Siemens timings, since the NIST recommended protocol of three repeats of 2D spin-echo sequence was found to be inaccurate and take a long scan duration (41.5 minutes).
[b]Philips NIST reference uses a 2D/SE T_2-mapping protocol with a composite broad band refocusing pulse rather than the sinc-shaped slice-selective refocusing pulse used in the UKRIN-MAPS MESE protocol.

using the harmonized UKRIN-MAPS MESE protocol and the NIST reference protocol. Reference T_2 values were temperature-corrected based on the recorded temperature using a linear regression model.[30]

To evaluate the accuracy of T_2 measurements, mean absolute percentage error (MAPE) was calculated by comparing the mean T_2 measurements from all pixels ($T_2(x,y)$) in spheres and repeats against reference values (T_2^{ref}):

$$\mathrm{MAPE} = \mathrm{mean}_{\mathrm{sphere}}\left(\frac{\left|\mathrm{mean}_{\mathrm{repeat}}\left(\mathrm{mean}_{\mathrm{pixel}}(T_2(x,y))\right) - T_2^{\mathrm{ref}}\right|}{T_2^{\mathrm{ref}}}\right) \times 100\%$$

The cumulative MAPE was calculated for those spheres with reference T_2 values in the physiologically relevant T_2 range (45–1286 msec).

In Vivo Experiments

Participants fasted for 2 hours prior to their scan session to limit dietary and hydration variability. As shown in Fig. 1, in vivo experiments consisted of two studies: 1) a "Travelling Kidney study" in which volunteers travelled and underwent scans at different imaging sites to assess intervendor variation; and 2) a "Repeatability study" in which volunteers were scanned multiple times at a single site.

The "Travelling Kidney study" was performed on eight healthy volunteers (five males/three females, age 32 ± 6 years (mean ± SD)), who were each scanned on all three vendors. For Siemens, the participants were scanned on either a Skyra Fit (five participants) or Prisma scanner (three participants) at two different imaging sites.

In the "Repeatability study," 10 healthy volunteers (five males/five females, age 32 ± 8 years (mean ± SD), five from the "Travelling Kidney" study group) were repeatedly scanned on a given scanner over a period of 2–6 months. Four participants were scanned two times on the Philips scanner, two participants were scanned four times on the GE scanner, and two participants were scanned four times on the Siemens scanners at two sites.

In both studies, the harmonized UKRIN-MAPS MESE protocol was used, which included the T_2 mapping acquisitions, B_1 mapping, T_2-weighted images, and MOLLI scans, the results of which are presented here. Whole kidney masks were automatically segmented from the T_2-weighted images using a convolutional neural network[28] (https://github.com/alexdaniel654/Renal_Segmentor). An operator (HL) with 10 years of experience in MRI manually segmented the cortex and medulla on the T_1 MOLLI maps, using an interactive graphical interface developed in MATLAB (R2019a, MathWorks Inc., Natick, MA).

The whole kidney, cortex, and medulla masks were applied to the T_2 maps with minor manual adjustments to correct for motion between acquisitions. This allowed for evaluation of the mean values and full-width-at-half-maximum (FWHM) local homogeneity of the T_2 distribution of voxels.

Fitting with Stimulated Echo Compensation

The "StimFit Toolbox,"[8,27] based on the EPG algorithm, was used to model stimulated-echo compensation in phantom and in vivo datasets. The EPG algorithm provides a system of equations that simulate the response to RF pulses with arbitrary flip angles including T_1 and T_2 relaxation effects.[14] Vendor-specific RF pulse shapes

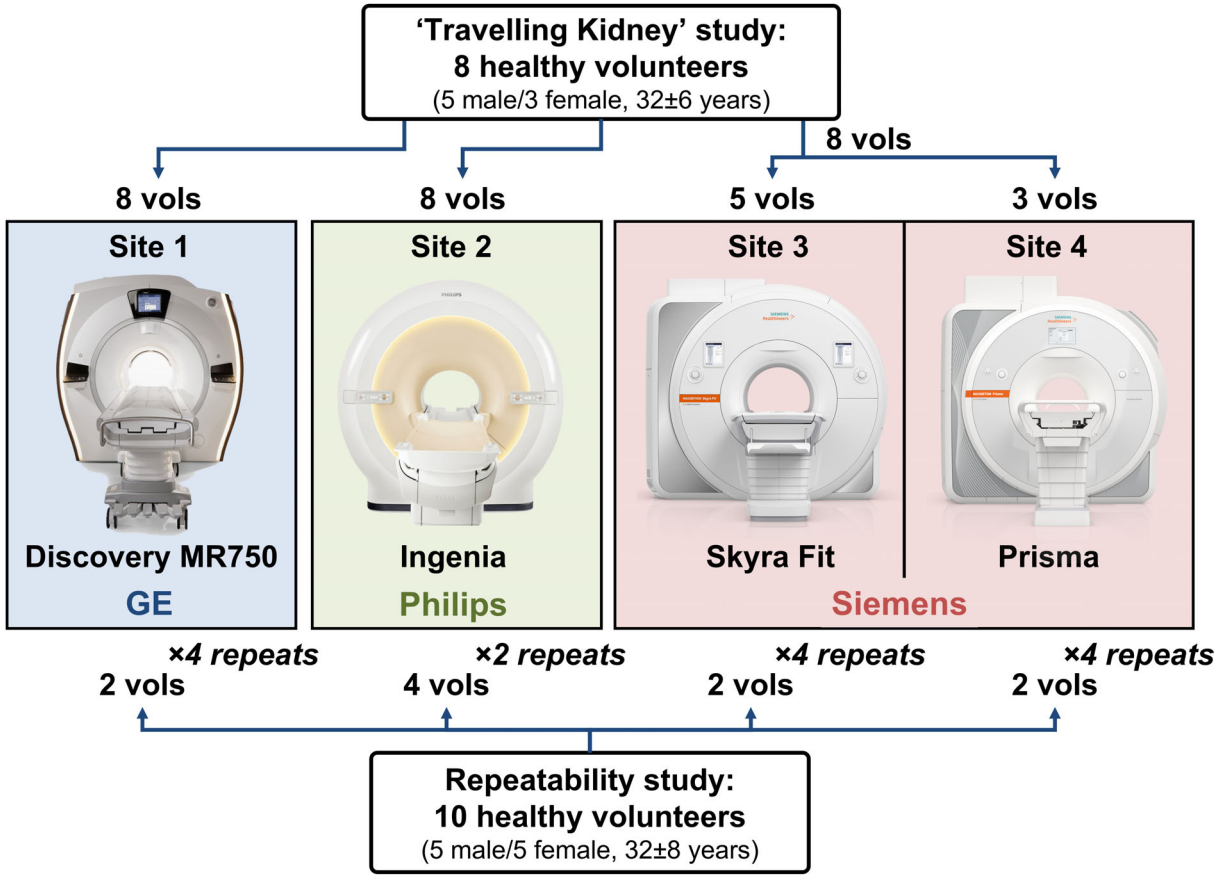


FIGURE 1: MR vendor, model, site information, and corresponding numbers of data sets (repeats) collected for healthy adult volunteers (vols) in the "Travelling Kidney" study and repeatability study.

and the nominal spatial width of excitation and refocusing pulses were input to "StimFit" to calculate the flip angle distributions across the slice profile, so that the effect of imperfect RF slice profiles could be accounted for. Magnetization evolving in alternate coherence pathways was assumed to experience negligible T_1 relaxation.[8] Furthermore, T_2 and B_1 values were estimated by a nonlinear least-squares algorithm with an objective function of an aggregate decay curve integrated over the slice profile.[27] Due to the symmetry of the spin-echo signal at refocusing angles surrounding 180°, "StimFit" precluded an estimated relative B_1 above unity (refocusing angle >180°) such that $0 \le B_1 \le 1$. For comparison, vendor-specific acquired B_1 maps were converted to the range [0, 1], i.e., converted B_1 (cB_1) = $(1 - \text{abs}\,(FA_{nominal} - FA_{actual})/FA_{nominal})$.

Statistical Analysis

Statistical analysis was performed in R software (version 4.2.2; https://www.r-project.org/) with packages "lme4" and "lmerTest." For the repeated phantom measurements, a random-intercept linear mixed-effects (LME) model was utilized to account for the data hierarchy. The data were entered as proportions relative to the temperature-corrected reference T_2 values. A categorical variable describing the fit method was studied as the fixed effect, and the intercept for specimen was modeled as the random effect. The *P*-values of the fixed effect were calculated using the Satterthwaite's degrees of freedom method.

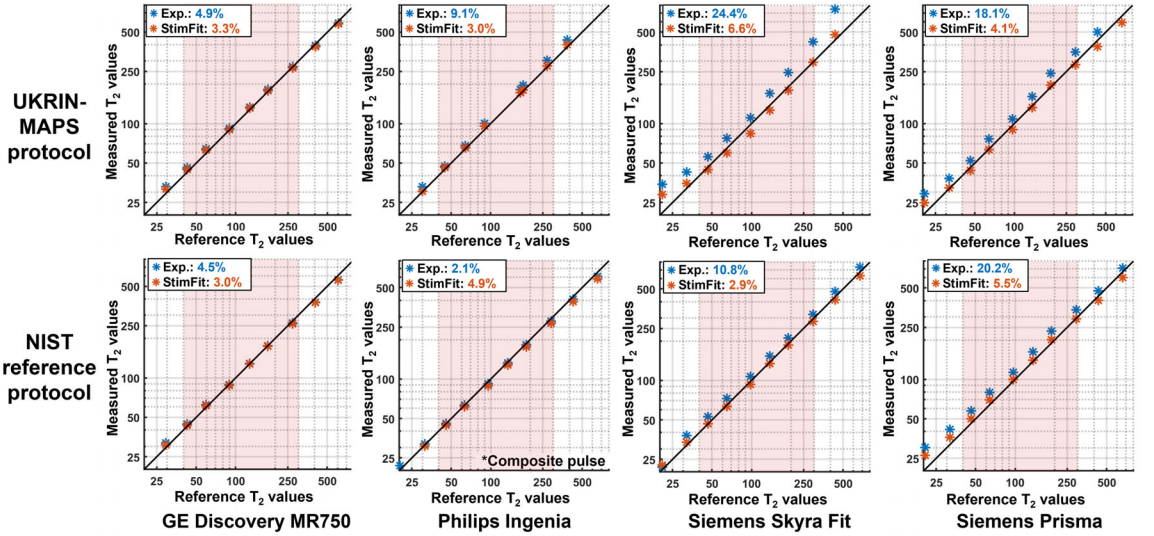
For in vivo measurements, a one-way analysis of variance (ANOVA) was performed to test for significant variations in T_2 measurements between vendors, and intervendor and intravendor coefficients of variation (CVs) were calculated. Paired student's *t*-tests were performed to compare T_2 mean and FWHM values, and CVs between the monoexponential fit and "StimFit." Unpaired student's *t*-tests were calculated to compare intervendor CVs (eight volunteers each scanned on three MR vendors) and intravendor CVs (10 volunteers each scanned by 1 MR vendor multiple times). Bland–Altman plots were generated to assess the consistency between each pair of vendors. *P*-values <0.05 were considered statistically significant for all analyses.

Results

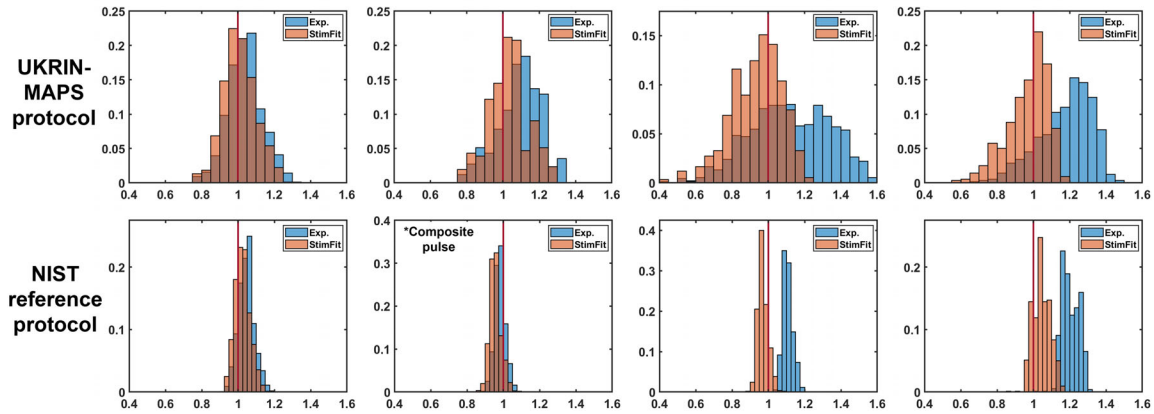
Phantom Experiments

Figure 2 and Table 2 show the T_2 measurements of the ISMRM/NIST system phantom across the different MR vendors and sites with reference values using a monoexponential fit and "StimFit" for both the UKRIN-MAPS and NIST reference protocol.

Compared with the monoexponential fit, "StimFit" reduced the MAPE of UKRIN-MAPS T_2 MESE measurements across the four sites (three vendors) from 4.9%, 9.1%, 24.4%, and 18.1% to 3.3%, 3.0%, 6.6%, and 4.1%, respectively.



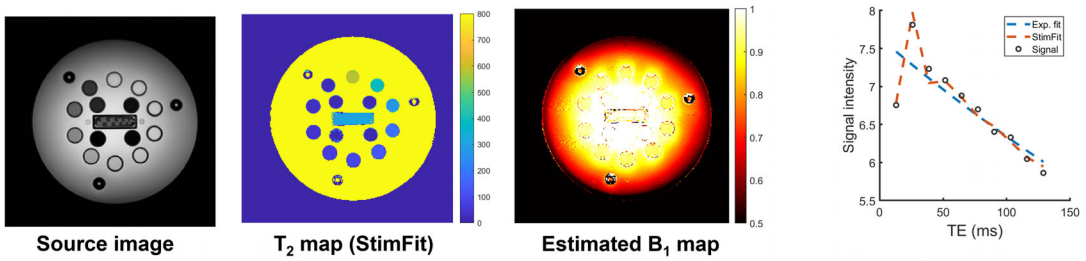
a T_2 measurements against reference values



b Histograms of normalized T_2 measurements

Vendor	Model	Temperature (℃)	Sphere Number									
			1	2	3	4	5	6	7	8	9	10
GE	Discovery MR750	17.5	612.9	408.4	273.3	177.3	128.8	90.0	59.7	43.0	29.7	19.2
Philips	Ingenia (NIST)	20.6	653.6	427.6	288.9	186.8	135.2	95.4	63.1	45.5	31.3	20.3
	Ingenia (UKRIN)	20.6	1928.3	379.5	267.3	175.0	166.9	88.9	63.4	44.2	29.9	19.4
Siemens	Skyra Fit	22	672.5	436.5	296.2	191.2	138.2	97.9	64.7	46.6	32.0	20.8
	Prisma	21.5	665.9	433.3	293.6	189.6	137.1	97.1	64.2	46.2	31.8	20.6
Reference		Base: 20	645.8	423.6	286	184.8	134.1	94.4	62.51	44.98	30.95	20.1

c Reference T_2 values (temperature-corrected)



d NIST phantom **e** Fitting curves

FIGURE 2: T_2 measurements computed using the exponential fit (Exp.) and "StimFit" from the ISMRM/NIST system phantom for different MR systems. Both the UKRIN-MAPS and NIST reference protocol were evaluated. (a) Average T_2 measurements (in msec) of each sphere against reference values (c) in the physiologically relevant range (45–286 msec, also indicated by the red rectangular boxes in a and c). The black boxes on the top show the MAPE. Apart from the Philips NIST protocol using composite pulses, "StimFit" reduced the MAPE of all measurements. (b) Histograms of T_2 measurements from all voxels within different spheres, normalized using corresponding NIST reference values. The red line in the center represents the baseline. (d) Example source image, T_2 map (in msec) and B_1 map from the phantom. (e) Example fitting curves of the exponential fit and "StimFit," noting the stimulated echo in the signal.

TABLE 2. T_2 Measurements from the ISMRM/NIST Phantom

	GE Discovery MR750	Philips Ingenia	Siemens Skyra Fit	Siemens Prisma
UKRIN-MAPS protocol				
MAPE				
Exponential fit	4.9%	9.1%	24.4%	18.1%
StimFit	3.3%	3.0%	6.6%	4.1%
LME *P*				
Exponential fit	0.04*	<0.001***	<0.001***	<0.001***
StimFit	0.82	0.007**	0.48	0.73
NIST protocol				
MAPE				
Exponential fit	4.5%	3.9%[a]	10.8%	20.2%
StimFit	3.0%	7.3%[a]	2.9%	5.5%
LME *P*				
Exponential fit	<0.001***	0.13	<0.001***	<0.001***
StimFit	0.1	<0.001***	0.34	0.08

The MAPE and the LME model were used to estimate the consistency of measurements with the reference values.
NIST = National Institute of Standards and Technology; UKRIN-MAPS = UK renal imaging network-MRI acquisition and processing standardization.
[a]Compared with the exponential fit, "StimFit" reduced the MAPE of all measurements except for the Philips NIST protocol, which uses composite pulses.
*$P < 0.05$;
**$P < 0.01$;
***$P < 0.001$.

For the NIST reference protocol, "StimFit" reduced as compared to an exponential fit the MAPE for the GE and the two Siemens scanners from 4.5%, 10.8%, and 20.2% to 3.0%, 2.9%, and 5.5%, respectively, while for the Philips NIST protocol that uses composite pulses a MAPE of 2.1% was found for the exponential fit as compared to 4.9% for StimFit.

Significant differences between T_2 measurements and reference T_2 values were found in all measurements using the monoexponential fit. In contrast, no significant difference was found between T_2 measurements and reference T_2 values for "StimFit" (Siemens Skyra Fit: NIST protocol $P = 0.34$, UKRIN protocol $P = 0.48$; Siemens Prisma: NIST protocol $P = 0.08$, UKRIN protocol $P = 0.73$; GE NIST protocol $P = 0.1$, UKRIN protocol $P = 0.82$) except for Philips which had a significant difference for the NIST protocol (composite pulses) and UKRIN protocol (with a small but consistent bias).

The correction of measurement bias by "StimFit" can also be observed in the histograms shown in Fig. 2b, which show the distribution of normalized T_2 measurements from all voxels across the different spheres in the T_2 array for the exponential fit and StimFit.

Example in Vivo Images

Figure 3 shows example monoexponential and "StimFit" in vivo T_2 maps, together with the estimated cB_1 maps from "StimFit" and the cB_1 maps computed from the separately acquired B_1 mapping sequences. The displayed maps are all from the same healthy volunteer collected across the three vendors. In regions where the flip angle was close to the nominal value, the T_2 maps agreed well between fitting methods and between vendors. However, flip angle variations due to non-ideal B_1 (cB_1 values with a discrepancy from 1) caused an overestimation in the monoexponential fit. This can be seen in the right kidney for data collected on the Philips scanner, and in the upper left kidney for the GE scanner, and in both kidneys for the Siemens dataset. For Siemens in particular, a widespread low B_1 caused a global overestimation of T_2 in all subjects. These overestimations in T_2 were largely corrected using "StimFit," which resulted in much more consistent T_2 values between kidneys and between vendors. Also,

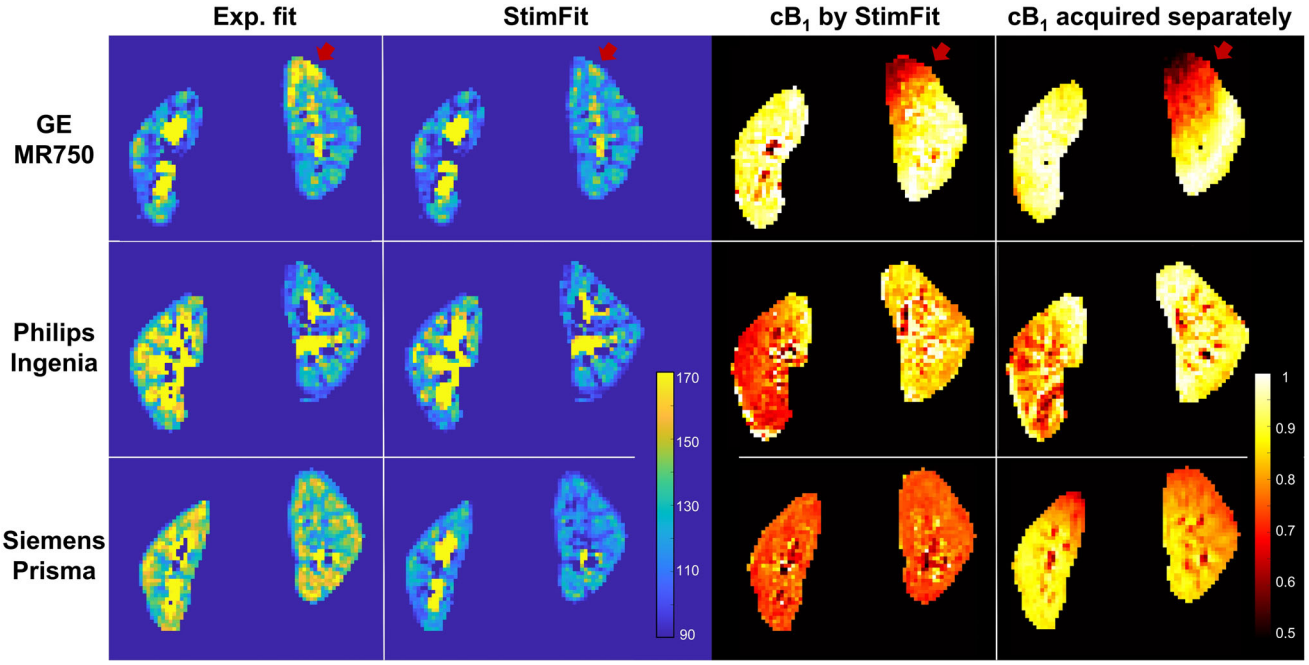


FIGURE 3: Example T_2 maps (in msec) processed by monoexponential fit (Exp. Fit) and StimFit from the same volunteer on the three MR vendors. Converted B_1 maps (cB_1) estimated by StimFit and acquired cB_1 maps are provided, these can be seen to show similar patterns and normalized intensities. Nonideal B_1 and corresponding overestimation by the monoexponential fit can be seen in the upper left kidney for data collected on GE (red arrows), in the right kidney for the Philips dataset, and in both kidneys for the Siemens dataset. These issues are corrected using "StimFit."

the cB_1 maps estimated by "StimFit" showed a similar pattern of features to the measured B_1 maps (but absolute values were not directly comparable due to the different RF pulses used in the T_2 mapping and B_1 mapping sequences).

Travelling Kidney Study

Bland–Altman plots in Fig. 4 show the T_2 measurements of the renal cortex and medulla from all volunteers between each two vendors. Specifically, "StimFit" corrected the bias of exponential fit results (cortex (msec): 2.6 vs. −1.6, −0.41 vs. 15, and −3 vs. 16; medulla (msec): 1.8 vs. −2.5, −1.2 vs. 12, and −3 vs. 15) and reduced the variance across vendors.

Figure 5 shows scatterplots of the T_2 measurements in the left and right whole kidneys from different vendors. Differences between left and right kidneys can be observed in the monoexponential fit results caused by B_1 inhomogeneity across the two kidneys. However, "StimFit" improved both the local T_2 homogeneity and T_2 variation across vendors.

The T_2 measurements in the cortex and medulla from eight healthy adult volunteers are summarized in Table 3 and Fig. 6a. The results from two Siemens scanners (Skyra Fit and Prisma) were combined due to their similar performance regarding measurements of the ISMRM/NIST phantom and similar B_1 field (average measured B_1 field of nominal flip angle: 82.3% vs. 82.7%). The T_2 measurements were significantly higher for monoexponential fit than "StimFit" in all vendors, particularly for Siemens with a widespread low B_1. For the monoexponential fit, significant differences were found in both cortex and medulla between vendors, whereas no significant difference was observed between vendors when using "StimFit" ($P = 0.86$ and $P = 0.92$). For the monoexponential fit, Siemens showed significantly higher T_2 measurements than the other two vendors, but "StimFit" results were consistent. The intervendor CVs were significantly reduced from 8.0% (cortex) and 7.1% (medulla) with exponential fit to 2.6% and 2.8% with "StimFit."

Repeatability Study: Intervendor and Intravendor Evaluation

Table 4 summarizes the measures of repeatability for mean T_2 values in the cortex and medulla. Specifically, "StimFit" reduced the intravendor CVs for most vendors compared to the monoexponential fit.

Figure 6b compares intervendor CVs from Table 3 with the intravendor CVs from Table 3. For "StimFit," there was no significant difference between the intervendor CVs (cortex: 2.61%, medulla: 2.76% in Table 3) and intravendor CVs (cortex: 1.26%, medulla: 1.62% in Table 4; $P = 0.13$ and $P = 0.05$). For the monoexponential fit, the intervendor CVs (cortex 8.0% and medulla 7.05% in Table 3) were significantly higher than the intravendor CVs (cortex 2.17% and medulla 2.5% in Table 4).

Assessing the T_2 Distribution in the Kidney

Figure 6c shows the FWHM of the T_2 distribution measured from cortex, medulla, and whole kidney. The FWHM was

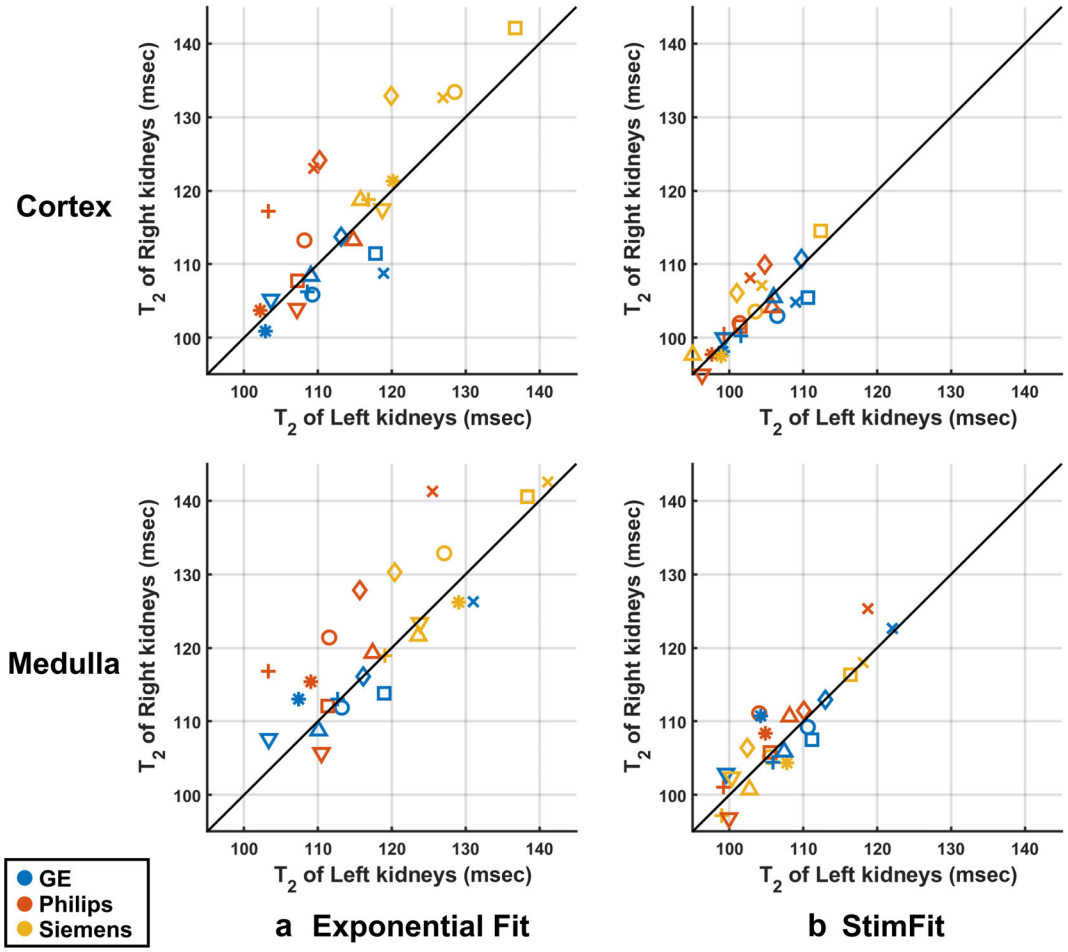


FIGURE 4: Bland–Altman plots showing the agreement between T_2 measurements of the monoexponential fit and "StimFit" methods between the different vendors. "StimFit" sufficiently reduced the variance across vendors. Each point corresponds to the measurement of one kidney.

significantly lower for "StimFit" compared to exponential fit in the cortex (StimFit: 6.9 ± 2.5, Exp. fit: 7.6 ± 2.8) and for the whole kidney (StimFit: 5.9 ± 2.5, Exp. fit: 6.3 ± 2.4).

Discussion

In this study, we demonstrated a large variance in renal T_2 mapping across MR vendors, despite using a harmonized MESE scheme with monoexponential fit on the same group of volunteers. By employing an EPG-based method (i.e., "StimFit"), the intervendor CVs were reduced to the same level as intravendor CV (<3%), so that no significant difference was found in the "StimFit" T_2 measurements between vendors.

It is worth noting that the monoexponential fit remains the default option on the vendor platforms evaluated in this study, and correction methods for T_2 mapping have not yet been recommended by the current consensus statements.[13,23]

When using the monoexponential fit, the measured T_2 values of cortex and medulla differed by up to 32 and 28 msec between vendors. This variation seems to be comparable to pathological T_2 changes reported in previous studies, such as 132 ± 22 msec and 97 ± 12 msec for high-grade and low-grade renal cell carcinomas,[3] and an increase from 77 ± 7 msec to 90 ± 6 msec after ischemia–reperfusion injury (in rabbits).[5] These findings suggest that the variability among vendors when using monoexponential fit may substantially impair the ability of T_2 mapping as a potential disease biomarker across multisite studies.

The inaccuracy and variance of measurements were mainly attributed to an imperfect B_1 field, which was revealed by both separate B_1 mapping acquisitions and the cB_1 maps estimated directly from the "StimFit" calculation. The B_1 field problems observed in this study included local B_1 inhomogeneities for the GE and Philips scanners, and overall B_1 miscalibrations for the Siemens scanners. Specifically, "StimFit" corrected these problems, resulting in accurate and homogeneous measurements. The improvement in local homogeneity of T_2 measurements was also demonstrated by a significant reduction in the FWHM for the cortex and whole kidney. Therefore, to address B_1 field problems, we recommend using EPG-based methods with B_1 correction instead of monoexponential fitting in multicenter studies. Additionally,

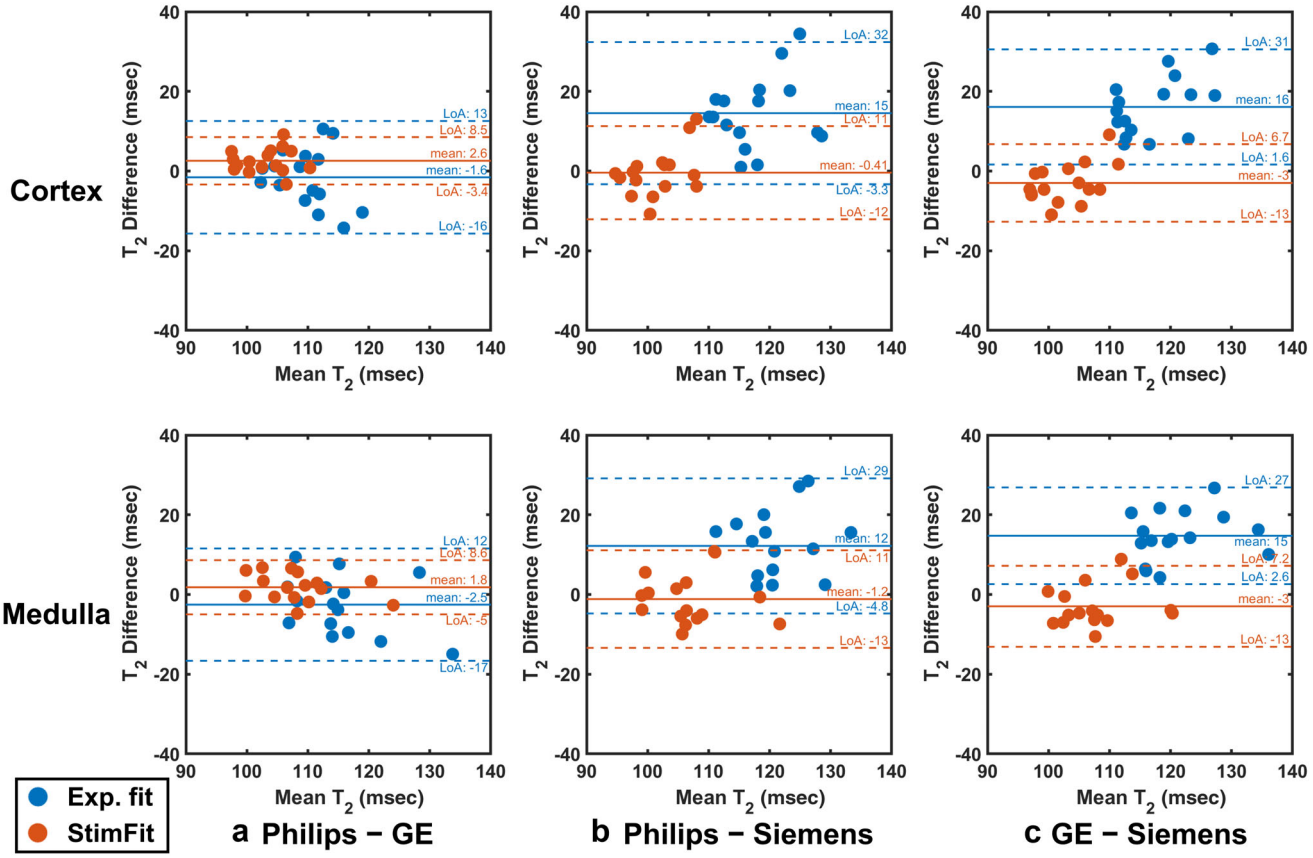


FIGURE 5: Scatterplots showing T_2 values in the left and right whole kidney obtained using the monoexponential fit (a) and "StimFit" (b). Different shapes correspond to the different subjects. "StimFit" can be seen to improve the local T_2 homogeneity and T_2 variation across vendors, including reducing the overestimation of the left kidney for Philips and the global overestimation for Siemens.

TABLE 3. Travelling Kidney Study: Mean T_2 Values in the Cortex and Medulla Obtained with the Monoexponential Fit and "StimFit" Averaged Over Eight Subjects for Different Vendors (msec, mean ± SD)

	GE	Philips	Siemens	Intervendor CV (%)	ANOVA (*P*)
Cortex					
Exp. fit	109.2 ± 4.8	110.6 ± 5.2	125.4 ± 8.0	8.00	<0.001
StimFit	104.4 ± 4.3	101.8 ± 4.1	101.7 ± 6.5	2.61	0.86
Difference	4.8 ± 1.8	8.7 ± 2.0	23.8 ± 2.1	–	–
Medulla					
Exp. fit	113.9 ± 7.0	116.3 ± 8.1	128.9 ± 8.0	7.05	0.002
StimFit	109.2 ± 6.4	107.5 ± 7.2	106.6 ± 7.1	2.76	0.92
Difference	4.6 ± 2.2	8.8 ± 2.1	22.3 ± 1.3	–	–
B_1					
Estimated cB_1 (%)	93.0 ± 2.1	82.7 ± 2.4	76.4 ± 1.2	–	–
Measured cB_1 (%)	89.3 ± 2.8	87.7 ± 3.4	82.5 ± 2.7	–	–

The coefficient of variance (CV) was calculated for T_2 values across different vendors. Difference = T_2 (Exp.) − T_2 (StimFit), which are significant in all measurements (paired *t*-test, $P < 0.001$).
ANOVA = analysis of variance; Exp. fit = exponential fit; cB_1 = converted B_1 (1 – abs ($FA_{nominal}$ − FA_{actual})/$FA_{nominal}$).

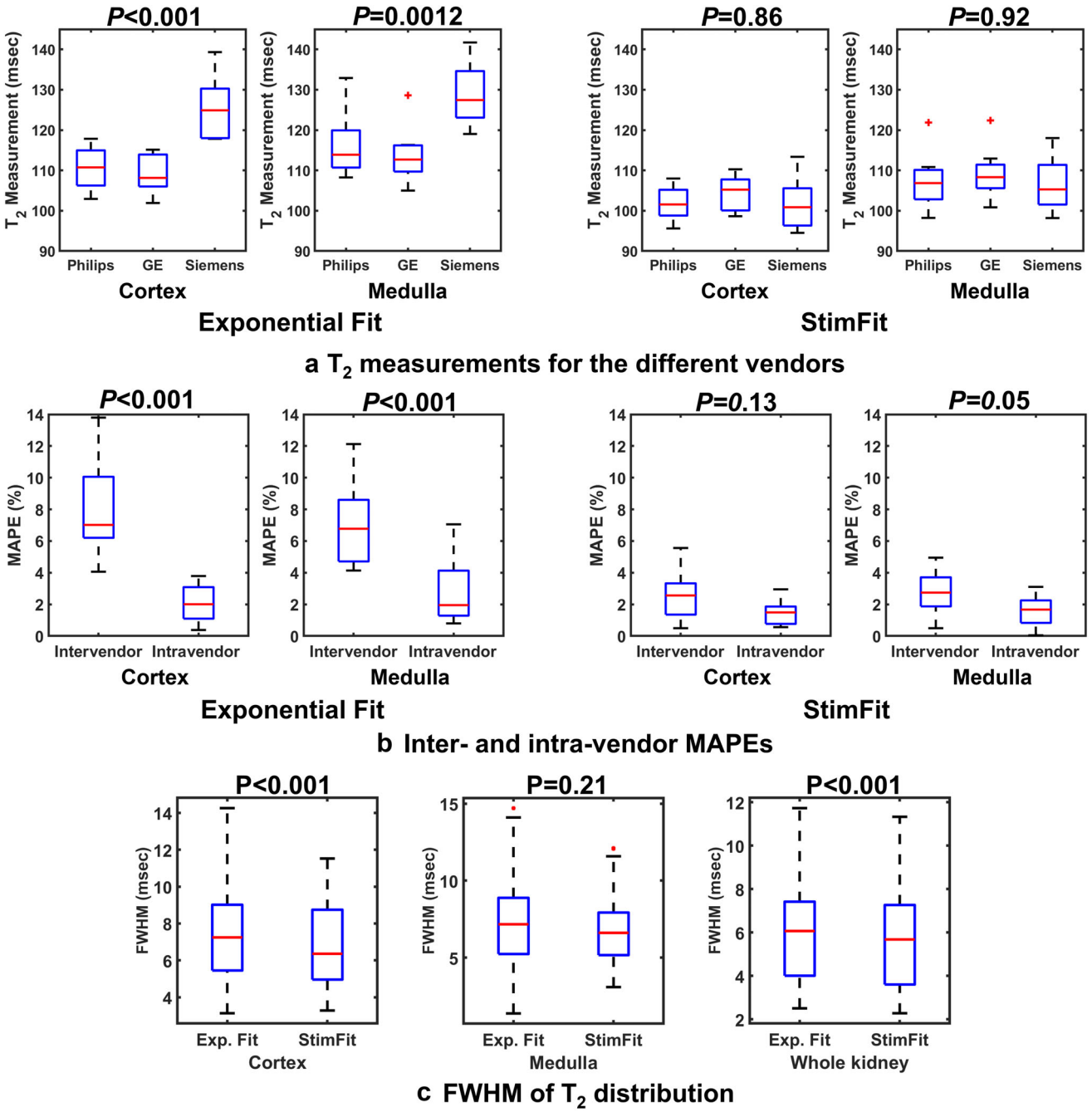


FIGURE 6: Boxplots comparing in vivo T_2 measurements using monoexponential fit and StimFit: (a) Mean T_2 values for different vendors. Significant differences in T_2 measurements were found between vendors for exponential fit, but not for StimFit (ANOVA). (b) Comparison of intervendor CVs from travelling volunteer scans and intravendor CVs from repeatability scans. Intervendor CVs were significantly higher than intravendor CVs for exponential fit, but not for StimFit (unpaired *t*-test). (c) The FWHM of the T_2 distribution measured from cortex, medulla, and whole kidney. The FWHM of T_2 measurements was significantly lower for StimFit compared to exponential fit in the cortex and whole kidney (paired *t*-test).

collecting a separate map of the transmit B_1 field to confirm this is advisable.

Stimulated echoes can be suppressed by composite rectangular pulses with optimized gradient crushers, which are less sensitive to changes in B_1. In this study, composite refocusing pulses were employed in the NIST reference protocol of Philips, which resulted in accurate T_2 measurements in the phantom (MAPE = 3.9% by exponential fit). However, it should be noted that composite pulses are not suitable for EPG models like "StimFit," as they destroy stimulated echoes due to the presence of large crushers[31]; hence, the lack of improvement when applying StimFit to the Philips NIST protocol as this uses composite pulses. Furthermore, compared with apodised sinc pulses, composite pulses cause a considerable increase in specific absorption rates.

TABLE 4. Repeatability Study: Intravendor CVs (%) of Repeating T_2 Measurements With Exponential Fit and "StimFit" for Repeatability Evaluation

	GE	Philips	Siemens Skyra Fit	Siemens Prisma	Average CV[a]
Number of repeats	2 vols × 4 repeats	4 vols × 2 repeats	2 vols × 4 repeats	2 vols × 4 repeats	10 vols
Cortex					
Exp. fit	2.47	1.50	2.56	2.84	2.17
StimFit	1.49	0.93	2.61	2.29	1.26
Medulla					
Exp. fit	2.69	2.04	2.41	3.31	2.50
StimFit	1.55	0.79	2.59	2.42	1.62

[a]The coefficient of variance (CV) was averaged over volunteers (vols) scanned by each scanner.

Main field (B_0) inhomogeneity effects were not addressed by the EPG model in StimFit, StimFit corrects only the T_2 inaccuracy due to the transmit field (B_1+) heterogeneity, with B_0 issues neglected in the models of the Bloch simulation. However, T_2 values have previously been shown to be robust to B_0 inhomogeneities, as well as variations in T_1 relaxation time and magnetization transfer.[18,20]

Limitations

A limitation of "StimFit" is that it requires the waveforms of excitation and refocusing pulses to be known, which are vendor-specific and may not be accessible for all scanners. Future studies will further need to investigate if a simpler and more general method can be effective for harmonization across vendors. Another limitation of this study is its small sample size, which only includes healthy subjects and does not investigate patients with relevant diseases. Future research will include groups of patients, including the planned 400 chronic kidney disease (CKD) patients collected in the AFiRM study (Application of Functional Renal MRI to improve assessment of CKD https://www.uhdb.nhs.uk/afirm-study/), to expand the investigation. In addition, we mainly focused on intervendor variations, but the possible variation between scanners within the same vendor has not been fully investigated. This study only included different scanners from one vendor (Siemens) at two different sites. The two scanners showed similar performance in T_2 measurements and B_1 homogeneities, and therefore their results were combined in the statistical analysis of in vivo results. More detailed evaluations are needed to investigate whether the interscanner variations originated from the differences between MR vendors or other configuration issues such as MR models, MR system versions, and transmit system types.

Conclusion

Variations in quantitative T_2 measurements in the kidney were observed across scanners and vendors despite using a harmonized MESE protocol, due to variability in the B_1 field. An EPG-based fitting method (i.e., "StimFit") reduces the B_1-associated errors and intervendor variations of measured renal T_2 values.

Acknowledgments

The authors acknowledge Dr. Ali Aghaeifar from Siemens Healthineers for providing RF pulse parameters.

Funding Information

This work was supported by Medical Research Council Grant (MR/R02264X/1 and UKRIN-MAPS), NIHR Nottingham Biomedical Research Centre (BRC-1215-200003), NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). Hao Li acknowledges the support from National Natural Science Foundation of China (No. 62201155) and the Shanghai Pujiang Program (No. 22PJ1400900).

References

1. Verhaert D, Thavendiranathan P, Giri S, et al. Direct T_2 quantification of myocardial edema in acute ischemic injury. JACC Cardiovasc Imaging 2011;4:269-278.
2. Franke M, Baeßler B, Vechtel J, et al. Magnetic resonance T_2 mapping and diffusion-weighted imaging for early detection of cystogenesis and response to therapy in a mouse model of polycystic kidney disease. Kidney Int 2017;92:1544-1554.
3. Adams LC, Bressem KK, Jurmeister P, et al. Use of quantitative T_2 mapping for the assessment of renal cell carcinomas: First results. Cancer Imaging 2019;19:1-11.
4. Hueper K, Rong S, Gutberlet M, et al. T_2 relaxation time and apparent diffusion coefficient for noninvasive assessment of renal pathology after acute kidney injury in mice: Comparison with histopathology. Invest Radiol 2013;48:834-842.

5. Chen J, Chen Q, Zhang J, et al. Value of T_2 mapping in the dynamic evaluation of renal ischemia-reperfusion injury. Acad Radiol 2022;29: 376-381.
6. Mathys C, Blondin D, Wittsack HJ, et al. T_2' imaging of native kidneys and renal allografts–a feasibility study. Rofo 2011;183:112-119.
7. Adams LC, Bressem KK, Scheibl S, et al. Multiparametric assessment of changes in renal tissue after kidney transplantation with quantitative MR relaxometry and diffusion-tensor imaging at 3 T. J Clin Med 2020; 9:1-16.
8. Lebel RM, Wilman AH. Transverse relaxometry with stimulated echo compensation. Magn Reson Med 2010;64:1005-1014.
9. Shridhar Konar A, Qian E, Geethanath S, et al. Quantitative imaging metrics derived from magnetic resonance fingerprinting using ISMRM/NIST MRI system phantom: An international multicenter repeatability and reproducibility study. Med Phys 2021;48:2438-2447.
10. Baeßler B, Schaarschmidt F, Stehning C, Schnackenburg B, Maintz D, Bunck AC. Cardiac T2-mapping using a fast gradient echo spin echo sequence–first in vitro and in vivo experience. J Cardiovasc Magn Reson 2015;17:67.
11. Huang TY, Liu YJ, Stemmer A, Poncelet BP. T_2 measurement of the human myocardium using a T2-prepared transient-state trueFISP sequence. Magn Reson Med 2007;57:960-966.
12. Deoni SCL, Rutt BK, Peters TM. Rapid combined T_1 and T_2 mapping using gradient recalled acquisition in the steady state. Magn Reson Med 2003;49:515-526.
13. Dekkers IA, de Boer A, Sharma K, et al. Consensus-based technical recommendations for clinical translation of renal T_1 and T_2 mapping MRI. Magn Reson Mater Phys Biol Med 2020;33:163-176.
14. Hennig J. Multiecho imaging sequences with low refocusing flip angles. J Magn Reson 1988;78:397-407.
15. Li H, Buchanan CE, Morris DM, et al. Improved harmonization of renal T_2 mapping between vendors using stimulated echo compensation. In: Proceeding Joint Annual Meeting ISMRM-ESMRMB ISMRT 31st Annual Meeting; 2022. 4409 p.
16. Kim D, Jensen JH, Wu EX, Sheth SS, Brittenham GM. Breathhold multiecho fast spin-echo pulse sequence for accurate R2 measurement in the heart and liver. Magn Reson Med 2009;62:300-306.
17. McPhee KC, Wilman AH. Transverse relaxation and flip angle mapping: Evaluation of simultaneous and independent methods using multiple spin echoes. Magn Reson Med 2017;77:2057-2065.
18. Ben-Eliezer N, Sodickson DK, Block KT. Rapid and accurate T_2 mapping from multi-spin-echo data using bloch-simulation-based reconstruction. Magn Reson Med 2015;73:809-817.
19. Huang C, Altbach MI, El Fakhri G. Pattern recognition for rapid T_2 mapping with stimulated echo compensation. Magn Reson Imaging 2014; 32:969-974.
20. Radunsky D, Stern N, Nassar J, Tsarfaty G, Blumenfeld-Katzir T, Ben-Eliezer N. Quantitative platform for accurate and reproducible assessment of transverse (T2) relaxation time. NMR Biomed 2021;34:e4537.
21. Buchanan C, Li H, Nery F, et al. Harmonisation of multiparametric renal MRI for multi-centre studies. In: 29th Annual Meeting of International Society for Magnetic Resonance in Medicine. 2021.
22. Buchanan CE, Li H, Morris DM, et al. A travelling kidney study using a harmonised multiparametric renal MRI protocol. In: Proceeding of 27th Annual Meeting ISMRM, Montr Canada; 2022. 482 p.
23. Mendichovszky I, Pullens P, Dekkers I, et al. Technical recommendations for clinical translation of renal MRI: A consensus project of the cooperation in science and technology action PARENCHIMA. Magn Reson Mater Phys Biol Med 2020;33:131-140.
24. Nehrke K, Börnert P. DREAM-a novel approach for robust, ultrafast, multislice B_1 mapping. Magn Reson Med 2012;68:1517-1526.
25. Klose U. Mapping of the radio frequency magnetic field with a MR snapshot flash technique. Med Phys 1992;19:1099-1104.
26. Sacolick LI, Wiesinger F, Hancu I, Vogel MW. B1 mapping by Bloch-Siegert shift. Magn Reson Med 2010;63:1315-1322.
27. Lebel RM: StimFit: A toolbox for robust T_2 mapping with stimulated echo compensation. In: Proceeding from 20th Annual Meeting ISMRM, Melbourne, Australia. 2012. 2558 p.
28. Daniel AJ, Buchanan CE, Allcock T, et al. Automated renal segmentation in healthy and chronic kidney disease subjects using a convolutional neural network. Magn Reson Med 2021;86:1125-1136.
29. Stupic KF, Ainslie M, Boss MA, et al. A standard system phantom for magnetic resonance imaging. Magn Reson Med 2021;86:1194-1211.
30. Statton BK, Smith J, Finnegan ME, Koerzdoerfer G, Quest RA, Grech-Sollars M. Temperature dependence, accuracy, and repeatability of T_1 and T_2 relaxation times for the ISMRM/NIST system phantom measured using MR fingerprinting. Magn Reson Med 2021;87:1446-1460.
31. Prasloski T, Mädler B, Xiang QS, MacKay A, Jones C. Applications of stimulated echo correction to multicomponent T_2 analysis. Magn Reson Med 2012;67:1803-1814.