Accuracy and repeatability of QRAPMASTER and MRF-vFA

Laura Nunez-Gonzalez a,*, Gyula Kotek a, Pedro A. Gómez b, Guido Buonincontri c, Mika Vogel d, Gabriel P. Krestin a, Dirk H.J. Poot a, Juan A. Hernandez-Tamames a

a Department of Radiology and Nuclear Medicine, Erasmus Medical Center, Rotterdam, the Netherlands
b Technische Universität München, Computer Science, Munich, Germany
c IMAGO7 Foundation, Pisa, Italy
d General Electric, GE Healthcare, Hoevelaken, the Netherlands

1. Introduction

Magnetic resonance imaging is widely used in clinical practice to detect and evaluate diseases. The signal level of each voxel depends not only on the nature of the tissue and the particular pulse sequence, but also on factors such as the strength of the main static magnetic field (B0) and its inhomogeneities, gradients and system imperfections such as B1 inacuracies. All of these confounding factors could contribute to a potential high variability of the weighted images that traditionally are acquired in follow-up and multi-site studies. In contrast to weighted images, quantitative MR relaxometry tries to minimize such variability by measuring actual relaxometry parameters. Several clinical studies have shown the relationship between quantitative maps and diseases such as multiple sclerosis, epilepsy, and dementia [1,2].

Despite of the advantages of quantitative MR relaxometry techniques and their potential capability to detect diseases, they are not yet routinely applied in clinical practice. There are two main reasons for this: long scanning time and concerns about the accuracy and repeatability of the quantitative MRI sequences QRAPMASTER, based on steady-state imaging, and variable flip angle MRF (MRF-vFA), based on the transient response. Both techniques are assessed with a standardized phantom and five volunteers on 1.5 T and 3 T clinical scanners. All scans were repeated eight times in consecutive weeks.

In the phantom, the mean bias ± 95% confidence interval for T1 values with QRAPMASTER was 10 ± 10% on 1.5 T and 4 ± 13% on 3.0 T. The mean bias for T1 values with MRF-vFA was 21 ± 17% on 1.5 T and 9 ± 9% on 3.0 T. For T2 values the mean bias with QRAPMASTER was 12 ± 5% on 1.5 T and 23 ± 1% on 3.0 T. For T2 values the mean bias with MRF-vFA was 17 ± 1% on 1.5 T and 19 ± 2% on 3.0 T. QRAPMASTER estimated lower T1 and T2 values than MRF-vFA. Repeatability was good with low coefficients of variation (CoV). Mean CoV ± 95% confidence interval for T1 were 3.2 ± 0.4% on 1.5 T and 4.5 ± 0.8% on 3.0 T with QRAPMASTER and 2.7% ± 0.2% on 1.5 T and 2.5 ± 0.2% on 3.0 T with MRF-vFA. For T2 were 3.3 ± 1.9% on 1.5 T and 3.2 ± 0.6% on 3.0 T with QRAPMASTER and 2.0 ± 0.4% on 1.5 T and 5.7 ± 1.0% on 3.0 T with MRF-vFA.

The in-vivo T1 and T2 are in the range of values previously reported by other authors. The in-vivo mean CoV ± 95% confidence interval in gray matter were for T1 1.7 ± 0.2% using QRAPMASTER and 0.7 ± 0.5% using MRF-vFA and for T2 were 0.9 ± 0.4% using QRAPMASTER and 2.4 ± 0.5% using MRF-vFA. In white matter were for T1 0.9 ± 0.3% using QRAPMASTER and 1.3 ± 1.1% using MRF-vFA and for T2 were 0.7 ± 0.4% using QRAPMASTER and 2.4 ± 0.4% using MRF-vFA. A GLM analysis showed that the variations in T1 and T2 mainly depend on the field strength and the subject, but not on the follow-up repetition in different days. This confirms the high repeatability of QRAPMASTER and MRF-vFA.

In summary, QRAPMASTER and MRF-vFA on both systems were highly repeatable with moderate accuracy, providing results comparable to standard references. While repeatability was similar for both methods, QRAPMASTER was more accurate. QRAPMASTER is a tested commercial product but MRF-vFA is 4.77 times faster, which would ease the inclusion of quantitative relaxometry.
repeatability of the quantifications. In fact, the long scan time has a negative impact on the clinical workflow. It hinders the incorporation of quantitative MRI into the clinical routine as well as adequate validation with fast conventional weighted imaging, such as SPGR, FIESTA, TSE/FSE, etc. To reduce the scan time, new fast quantitative multi-parametric imaging methods have been developed. Some of them based on steady-state pulse sequences, such as “Quantification of Relaxation Times and Proton Density by Multi-Echo acquisition of a saturation-recovery using Turbo spin-Echo Readout” (QRAPMASTER) [3] or PLANET [4]. Some others have been recently developed based on the transient response such as Magnetic Resonance Fingerprinting (MRF) [5], and MRF variable flip angle (MRF-vFA) [6]. Unlike other quantitative imaging methods that are commonly accepted such as DESPOT1 and DESPOT2 [7], those techniques obtain several quantitative maps within a single scan.

Despite the clear potential of these new techniques, to be used in research or clinical practice they have to demonstrate good accuracy and precision. The appropriate evaluation of the consistency of quantitative values is itself a challenge that requires evaluation of the bias and the repeatability [8]. Consistently with the metrology methods described in [8], clinically acceptable in-vivo bias and repeatability have been measured using QRAPMASTER [9–15]. Also, the bias and repeatability of the T1 and T2 quantification have been studied with other methods relying on the steady-state [7,9,16]. In this work we compare bias and repeatability of MRF-vFA and QRAPMASTER with a test-retest study using a standardized phantom and healthy volunteers. To study the applicability for clinical workflow, both 1.5 T and 3.0 T clinical scanners are used.

2. Methods

In this work we evaluated the accuracy and repeatability of two techniques using Bland-Altman plots [19], coefficients of variation (CoV) [18] and a General Linear Model (GLM) analysis [20]. Both techniques, QRAPMASTER and MRF-vFA, were provided by the systems vendor, General Electric Medical Systems. QRAPMASTER as a final product and MRF-vFA as research sequence.

2.1. Description of the techniques

QRAPMASTER uses a multi-echo acquisition of a saturation-recovery sequence with Turbo Spin-Echo (TSE) Readout to obtain quantitative maps of T1, T2, and Proton Density (PD) [3]. For each slice a saturation pulse is applied and after some delay a series of excitation pulses of 90 degrees and refocusing pulses of 180 are played to obtain a multi spin-echo acquisition. This multi spin-echo acquisition allows quantifying the T2 values. By repeating the same sequence for the same slice with different delays the T1 values can be calculated. Finally, with the information of T1 and T2, the PD is calculated from the intensity at echo time zero.

Once the quantitative maps have been obtained, T1 and T2 weighted images are synthesized from these maps. This sequence was acquired using Magnetic Resonance Image Compilation (MAGIC), which is a customized version of the package SyMRI IMAGE [21].

In contrast to QRAPMASTER, conventional MRF and MRF-vFA use a single rapid acquisition to retrieve a signal evolution that is sensitive to T1 and T2 during the transient response [5], not relying on steady-state models for parameter inference. Along the acquisition, the readout trajectory and some acquisition parameters such as repetition time (TR), echo time (TE) and pulse phase can be varied. Subsequently, the relaxation parameters are recovered by matching the signal evolution to a dictionary that, for many combinations of T1 and T2, can be pre-computed using the extended phase graph formalism (EPG) [6,22].

In contrast to conventional MRF, MRF-vFA uses a constant TR and the flip angles are linearly increasing along the acquisition. Therefore, in the case of the MRF-vFA for each slice, first an inversion pulse is played out as in MRF. Then a series of excitation pulses is applied with constant TR and TE. Additionally, the flip angle linearly increases from low flip angles (5 degrees) to high flip angles (70 degrees). After each excitation pulse, one arm from the spiral trajectory is acquired being rotated by a golden angle after each excitation pulse. Small flip angles after the inversion pulse allow capturing the recovery of the longitudinal magnetization for T1 estimation and the increase in the flip angle procure differentiation of the signal evolution for different T2 values. Furthermore, in order to improve estimations with a low number of TRs, MRF-vFA applies a compressed sensing reconstruction [6] to obtain unalised images before parameter inference. The main advantage of using the MRF-vFA acquisition scheme [6] is that it allows at least 3 times reduction of the acquisition time compared to a zero-filled MRF reconstruction with the original MRF scheme [6].

2.2. Acquisition

The acquisitions for the Institutional Review Board-approved study were performed on two systems, a 1.5 T GE MR450 and a 3.0 T GE MR750 (General Electric Medical Systems, Waukesha, WI). In both systems, a 16 channel Head, Neck and Spine array coil was used.

For QRAPMASTER the acquisition parameters are given in Table 1. The scan time to acquire 27 slices was 5 min and 34 s. MRF-vFA used 260 TRs with constant TR = 10 ms (Table 1). The flip angle was increased continually from 5 to 70 degrees as described in [6]. The scan time for 27 slices was 1 min and 10 s.

Axial images from the NIST/ISMRM System Phantom [23] were acquired. To minimize temperature induced variations the phantom was kept in the scanner room of the 3.0 T system. The temperature inside the 3.0 T scanner room was measured to be between 22.6 °C and 22.9 °C. The temperature inside the 1.5 T system room was measured to be between 22.5 °C and 23 °C.

A sampled size of 5 was calculated using G*Power [24] with a specificity of 95%, sensitivity of 90% and a correlation of 90% among repeated measurements, obtaining an statistical power of 0.92. This study design was approved by the Institutional Review Board. After providing informed consent, five healthy volunteers (3 females and 2 males, between 18 and 25 years) participated in the repeatability study. The READYBrain sequence [25] was used to align each quantitative MRI acquisition to the AC-PC plane. READYBrain automatically detects the AC-PC plane for each subject, facilitating the registration and the segmentation.

To assess the repeatability of QRAPMASTER and MRF-vFA the five different acquisitions for QRAPMASTER and MRF-vFA for the phantom and the in-vivo (phantom/in-vivo when the parameter differs).

| Technique | QRAPMASTER | MRF-vFA |
|-----------|------------|---------|
| SequenceType | 2D - Turbo Spin Echo | 2D - SpoiledGradient Echo |
| Preparation | Saturation pulse | Inversion pulse |
| K-space trajectory | Cartesian, full k-space | Spiral, under-sampled k-space |
| Orientation | Axial/AC-PC | Axial/AC-PC |
| FOV (cm) | 31 | 31 |
| Voxel size (mm × mm × mm) | 1.2 × 1.2 × 5.0 | 1.2 × 1.2 × 5.0 |
| Echo Train Length/# TRs | 12 | 260 |
| Flip Angle (degrees) | 90 | 5.70 |
| TR (ms) | 4700/4400 | 10 |
| Acquisition time (27 slices) | 5 min 34 s | 1 min 10 s |
volunteers and the phantom were scanned with the protocol mentioned above in both systems on the same week-day for 8 consecutive weeks.

2.3. Reconstruction and parameter estimation

The data acquired with MRF-vFA were reconstructed using the non-uniform Fast Fourier algorithm [26] and temporal subspace reconstruction. The used solver is the alternating direction method of multipliers (ADMM) with 10 basis coefficients and regularized with local low rank regularization on spatiotemporal patches of dimension $[8 \times 8 \times 10]$ [27].

For MRF-vFA, slice-profile imperfections were included in the dictionary generated with the extended phase graph formalism [22,28]. We calculated a specific slice profile for each field strength. Therefore, separate dictionaries were created for the 1.5 T and 3.0 T systems. For both dictionaries, the range of T1 values was from 100 ms to 3000 ms with steps of 10 ms and the range of T2 values was from 10 ms to 1000 ms with steps of 5 ms. The selected T1 and T2 ranges were typical for brain tissue [7,29,30].

After the images were reconstructed, the signal evolution for each voxel was matched to the dictionary and the best match, i.e. the atom with the highest correlation, provided the estimate for the PD, T1 and T2 values.

In the case of QRAPMASTER, only synthetic T1-weighted, T2-weighted and FLAIR images provided by MAGiC could be exported with any repetition time (TR), echo time (TE) and inversion time (TI). Images with different TR and TE were exported and used to obtain the PD, T1 and T2 maps. The study of the PD, with T1 and T2, was performed only for in-vivo experiments and it was excluded from the phantom analysis.

2.4. Data preparation

Before the analysis, all the images were converted from DICOM (Digital Imaging and Communication On Medicine [31]) to NIFTI using the Statistical Parametric Mapping (SPM12) toolbox for MATLAB (Mathworks, Natick MA) [32].

For the phantom, a 2D plane through the center of the T2 contrast spheres was selected. We chose those spheres within the range supported by QRAPMASTER ($T1 \geq 300$ ms and $T2 \leq 250$ ms). A region of interest (ROI) was drawn for each one (Fig. 1) and the average inside the ROIs was calculated. Perfect spatial alignment was assumed between maps acquired with different techniques in the same scan session. The T1-weighted image created with QRAPMASTER images from separate

|          | 1     | 2     | 3     | 4     | 5     | 6     |
|----------|-------|-------|-------|-------|-------|-------|
| 1.5 T    |       |       |       |       |       |       |
| T1 (ms)  | 1237.00 (0.40) | 1030.00 (1.70) | 752.20 (1.12) | 550.20 (0.18) | 413.40 (0.29) | 292.90 (0.15) |
| T2 (ms)  | 184.90 (0.110) | 140.60 (0.050) | 91.76 (0.029) | 64.84 (0.029) | 45.28 (0.029) | 30.62 (0.014) |
| 3.0 T    |       |       |       |       |       |       |
| T1 (ms)  | 1332.00 (0.80) | 1044.00 (3.20) | 801.70 (1.70) | 608.60 (1.03) | 458.40 (0.33) | 336.50 (0.18) |
| T2 (ms)  | 133.27 (0.073) | 96.89 (0.049) | 64.07 (0.034) | 46.42 (0.014) | 31.97 (0.083) | 22.56 (0.012) |

Fig. 1. Synthetic T1-weighted image of plate 4 of the NIST phantom. The ROIs are outlined in green and the nominal values for these ROIs are shown in the table. The standard deviation is reported in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
sessions were all rigidly co-registered [33] to the first scan session using SPM12 [32] allowing us to use the same ROIs for all the images.

For the in-vivo data, the pipeline was slightly different. First, co-registration within the same scan-session was needed to correct motion-induced misalignment of the images for both MRF-vFA and QRAPMASTER. For each session, all images were rigidly co-registered to a synthetic T1-weighted image created with QRAPMASTER. This T1-weighted image was also used to segment gray matter and white matter with SPM12 [32,34]. Once the segmentation was completed, the results were applied to compute the mean values inside the gray and white matter of the quantitative maps within the same subject and scan-session. The mean values of all the subjects for gray matter and white matter were calculated for each technique, system, and scan-session.

The last procedure applied to the in-vivo data was the creation of global maps using all the brain acquisitions. A T1-weighted template registered to the MNI (Montreal Neurological Institute) space [35], which defines a standard brain by using a large series of MRI scans on normal controls, was created using Diffeomorphic Anatomical Registration through Exponentiated Lie algebra (DARTEL) algorithm [36] with SPM12, including all the T1-weighted images created with QRAPMASTER in both scanners. The template generation was used to normalize the acquisitions to MNI space. It has been shown that diffeomorphic registration processing could bias the results [37–39] obtained in longitudinal studies because of the asymmetries of the non-linear interpolations applied. To reduce the impact of it, all the acquisitions were treated equally by creating a template with all the acquisitions (forward) and then aligning the acquisitions to this template (backward) [38]. The normalized acquisitions were used both to study the voxel-wise repeatability and as the input of for a GLM analysis as shown in the Supplementary Fig. S1.

The relative proton density maps were self-normalized (nPD) to the average inside the brain mask in each case as in [17], in order to avoid variations due to coil sensitivity and scaling.

2.5. Data analysis

2.5.1. Phantom

Accuracy of phantom acquisitions was assessed by comparing the ROI means to the nominal values as reported in the specifications manual [23]. The variation of the measurements reported by the manufacturer is negligible since they are in the order of 1 ms or less except for the T1 value of the second ROI on 3.0 T (standard deviation is 3.2 ms). The agreement between the ROI mean of each session and the nominal values was assessed through Bland-Altman plots [19].

To assess systematic differences in the estimated values using MRF-vFA and QRAPMASTER, we performed linear regression between their estimated ROI mean T1 and T2 values of the phantom. The linear regression was performed with Total Least Squares [40]. It provided the proportionality (slope) between the estimated values of MRF-vFA and QRAPMASTER as well as the offset (intercept). Additionally Pearson’s coefficient of correlation (R-value) between the ROI mean T1 and T2 values of MRF-vFA and QRAPMASTER was calculated.

The repeatability was quantified with the coefficient of variation (CoV = \(\frac{\sigma}{\mu}\)) [18] where \(\sigma\) is the standard deviation and \(\mu\) is the mean of the averaged values within the ROIs over the 8 scan sessions and its 95% confidence interval of the CoV was calculated [41].

2.5.2. In-vivo

The accuracy was evaluated by comparison to the values reported by previous studies [7,13,21,29,42,43]. The relative bias for each technique, was calculated using the median of the reported values using the inversion-recovery (T1) and CPGM (T2) standard relaxometry methods.

For each volunteer the in-vivo repeatability was assessed similarly as for the phantom, using the per session gray and white matter segmentation as the ROIs to evaluate. For each subject, the CoVs of gray and white matter were obtained from the average values over the segmented tissues. These CoVs were averaged over all the subjects (averaged-tissue-CoVs). Separately, voxel-wise CoV maps were obtained in MNI space for each subject. These individual CoV maps were averaged over the volunteers into mean-CoV maps to study the dependence of the variability with the anatomy.

Besides the CoVs, a GLM analysis [20] was used to assess the repeatability of the quantitative maps nPD, T1 and T2 as in [17] using the implementation provided by SPM12 [32]. In our case, we modeled as binary covariates (b) the subjects (1–5), the field strength (1.5 T and 3.0 T), and the week of acquisition (1–8). Supplementary Fig. S1 shows the design matrix that is used separately for each method and for each voxel of the nPD, T1 and T2 maps, normalized to MNI space.

The estimated bmaps from the GLM quantify the influence of the corresponding covariate on the nPD, T1 and T2 maps. To analyze how the covariates influence the variability among the quantitative maps, we calculated the average of the bmaps over tissue maps (normalize to the MNI space), while to take into account possible regional effects of variability; we extracted the bmaps for a representative slice.

3. Results

3.1. Phantom

Fig. 1 shows the selected slice of the phantom as well as the ROI’s and their nominal T1 and T2 values with the standard deviation in brackets. Fig. 2 shows that the estimated T1 and T2 values for both techniques were within the limits of agreement respect to the reference values according to the Bland-Altman analysis. Table 2 shows the mean bias from the Bland-Altman analysis and the 95% limits of agreement. The bias was positive in all cases, including the null value only for T2 estimations on the 1.5 T system and T1 estimations on the 3.0 T system using QRAPMASTER, and T2 estimations on the 1.5 T system using MRF-vFA. This means that the techniques are overestimating the values, as it can be seen in Fig. 2. Also, it can be appreciated that the mean bias with MRF-vFA was double than with QRAPMASTER for T1 values. However, for T2 values, the mean bias was high and similar for QRAPMASTER and MRF-vFA. In general, the spread for different days was similar across the values, except in the case of T2 values estimated on the 1.5 T system, where lower T2 values had less dispersion. However, they had a higher bias than longer T2 values, which means a constant bias in milliseconds penalizing more the relative bias for smaller T2 values.

Table 2 and Fig. 3 show the results of the total least squares regression between the average ROI mean T1 and T2 values obtained by MRF-vFA and QRAPMASTER. The methods are strongly correlated (R-values > 0.99). The slopes for the T1 and T2 estimated values on the for both 1.5 T system and on the 3.0 T system were larger than one. This means that for MRF-vFA the spread in T1 and T2 values over the ROIs is larger than for QRAPMASTER. Consequently, the intercept is negative, except for the estimate T1 values on the 3.0 T system where an intercept of zero is included within the 95% CI.

Fig. 4 shows the Bland-Altman plots between MRF-vFA and QRAPMASTER. In general, the estimated values with MRF-vFA are within the limits of agreement comparing to QRAPMASTER. Only few sessions were outliers in which the estimated values of T1 were outside the limits of agreement on the 1.5 T system and on the 3.0 T system. In the case of the T2 values, only on the 1.5 T system some values estimated with MRF-vFA fell outside the limits of agreement.

Fig. 5 assesses the repeatability by showing the CoV of the ROI means. In all the cases, the CoV was less than 8%. Note that the CoV is smaller than the bias, especially for T2 values on the 3.0 T system, which present high bias but had low CoVs. Also, both techniques showed higher CoVs on the 3.0 T system than on the 1.5 T system. However, the performance of QRAPMASTER and MRF-vFA was different depending on the relaxation parameter (T1 or T2) and the system.
On the 3.0 T system QRAPMASTER was more variable for T1 while MRF-vFA was more variable for T2 (the CoVs were 1%–4% higher for the same ROIs).

On 1.5 T, QRAPMASTER showed CoVs 1% higher for longer T1s (T1 > 1000 ms) and CoVs 2–3 times higher in the case of long T2 values (T2 > 100 ms). In contrast to the results on 3.0 T, the CoVs for the estimated T1 and T2 values using MRF-vFA were similar.

On both systems, the variability of QRAPMASTER was more dependent on the T1s and T2s in the phantom than MRF-vFA, since the CoV was higher for longer T1 and T2 values (T1 > 1000 ms and T2 > 100 ms, ROI 1 and 2).

### 3.2. In-vivo

The quantitative T1 and T2 maps provided by both techniques showed a good contrast (Fig. 6). However, QRAPMASTER showed T1 maps with higher contrast between gray and white matter than MRF-vFA, while MRF-vFA showed T2 maps with higher contrast than QRAPMASTER. This can be observed in multiple slices of the Supplementary Figs. S4–S15. Also these differences in the contrast are present in the synthetic images created from the PD, T1 and T2 maps, as in the example showed in Supplementary Figs. S16 and S17.

Table 3 shows the average value and the standard deviation of the computed-tissue-values of the gray and white matter over the 8 sessions. In all the cases, the estimated T1 and T2 values were in the range of the values reported in previous studies. The average T1 value in gray matter...
Table 2
In the first two columns, the mean biases and the 95% limits of agreement (LOA) from the Bland-Altman analysis are reported for QRAPMASTER and MRF-vFA.

| Field strength | Relaxation parameter | QRAPMASTER (95% CI) [95% LOA] (95% CI) | MRF-vFA (95% CI) [95% LOA] (95% CI) | Slope | Intercept (ms) | R-value |
|----------------|----------------------|----------------------------------------|---------------------------------------|--------|----------------|----------|
| 1.5 T          | T1                   | 10.3 (±9.65) [-5.1, 25.7] (±3.9)       | 20.8 (±17.22) [11.1, 30.5] (±2.5)     | 1.157  | −36.298        | 0.999    |
|                | T2                   | 12.4 (±3.09) [-17.2, 41.9] (±7.5)     | 16.8 (±1.11) [-5.3, 38.9] (±5.6)     | 1.164  | −9.529         | 0.999    |
| 3.0 T          | T1                   | 3.9 (±13.33) [-5.5, 13.6] (±2.4)      | 9.5 (±9.26) [3.5, 15.6] (±1.5)        | 1.051  | 1.091          | 0.999    |
|                | T2                   | 22.7 (±1.15) [1.5, 43.9] (±5.4)       | 19.2 (±2.04) [2.5, 35.9] (±4.2)       | 1.108  | −8.417         | 0.993    |

Between brackets the 95% confidence interval of the bias and the LOA. In the following columns, the results of the linear regression between the estimated values with MRF-vFA and the values estimated with QRAPMASTER. The values of the slopes, intercept and correlation (R-value) resulted from the linear regression between both techniques are shown. The 95% confidence interval for the slope and the intercept is presented in brackets.

Fig. 3. The estimated T1(A, B) and T2 (C, D) values obtained from the NIST phantom by MRF-vFA against the values obtained by QRAPMASTER on the 1.5 T system (A, C) and the 3.0 T system (B, D). Vertical and horizontal lines mark the standard deviation for MRF-vFA and QRAPMASTER, respectively. A solid black line marks the fitted line. The dashed black lines mark the 95% confidence interval boundaries. The blue line is the identity line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
was lower for MRF-vFA than for QRAPMASTER, while in the case of white matter, QRAPMASTER produced lower T1 values than with MRF-vFA. In the case of T2, QRAPMASTER estimated larger values for gray and white matter than MRF-vFA. While QRAPMASTER estimated around 30% higher nPD for gray matter than for white matter, for MRF-vFA the difference in estimated nPD for gray and white matter was smaller than 2%.

Table 4 reports the median values found in literature and the bias of the averaged values for QRAPMASTER and MRF-vFA to these values. In contrast to the phantom results, the bias was higher on the 3.0 T system. Despite QRAPMASTER had lower bias than MRF-vFA in the phantom experiment, in gray matter the bias of QRAPMASTER was larger than that of MRF-vFA, except when estimating T1 values on the 3.0 T system. Furthermore, the bias estimating T1 in white matter was higher using QRAPMASTER than MRF-vFA. Also, the T1 was underestimated in contrast to the positive bias found on the phantom, except for the gray matter using QRAPMASTER on the 1.5 T system.

Fig. 7 shows the averaged-tissue-CoVs (in percentage) on the 1.5 T and the 3.0 T system. All the CoVs were below 4%. On the 1.5 T system (Fig. 7-A), the CoV with QRAPMASTER were lower than with MRF-vFA, except for T1 estimated values in gray matter. On the 3.0 T system (Fig. 7-B), the CoVs for MRF-vFA were lower than for QRAPMASTER for T1 values, while for T2 values, QRAPMASTER had lower CoVs than MRF-vFA. Using QRAPMASTER, the CoVs were higher for T1 estimated values than for T2 estimated values, on both systems. In contrast, using MRF-vFA the CoVs were higher for T2 estimated values than for T1 estimated values.

We observed similar range dependence for the QRAPMASTER in the CoVs for the in-vivo data (Fig. 7) than in the phantom. The estimated T1 values for gray matter were above 1000 ms (Table 3) and the CoVs for gray matter are higher than for white matter on both systems.

Fig. 8 shows the mean-CoV maps (in percentage). The mean-CoVs were similar on both systems, but they differed between the techniques as well as among relaxation parameters. QRAPMASTER showed a lower mean CoV for the nPD and T2 values than MRF-vFA, while the estimation of T1 values with QRAPMASTER had higher mean-CoVs than with MRF-vFA. We observed higher mean-CoV for the CSF and the skull area than in gray and white matter. In the rest of the brain, the mean-CoV was low and similar between different areas.
Fig. 5. The coefficient of variation of T1 (top) and T2 (bottom) on a 1.5 T system (left) and on a 3.0 T system (right) from the NIST phantom using QRAPMASTER (blue) and MRF-vFA (yellow). Vertical lines show the 95% confidence intervals in blue (QRAPMASTER) and red (MRF-vFA). The ROIs are shown in Fig. 1. Horizontal axis labels show ROI numbers and corresponding T1 values (A, B) and T2 values (B, D) are in brackets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Quantitative PD (left), T1 (middle), and T2 (right) maps from a single subject obtained by QRAPMASTER (top) and MRF-vFA (bottom) on a 1.5 T system (odd rows) and on a 3.0 T system (even rows).
Table 3
Estimated relaxation parameters from in-vivo data.

| Tissue      | Relaxation parameter | 1.5 T                   | 3.0 T                   |
|-------------|----------------------|-------------------------|-------------------------|
|             |                      | QRAPMASTER (ms)         | MRF-vFA (ms)            |
| Gray Matter | T1                   | 1057 (± 28)             | 964 (± 17)              |
|             | T2                   | 96 (± 2)                | 80 (± 2)                |
|             | nPD                  | 1.07 (± 0.011)          | 0.99 (± 0.012)          |
| White Matter| T1                   | 501 (± 15)              | 699 (± 25)              |
|             | T2                   | 83 (± 2)                | 61 (± 2)                |
|             | nPD                  | 0.83 (± 0.016)          | 0.97 (± 0.016)          |

Table 4
Median of the T1 and T2 values (range) for gray matter found in literature for inversion recovery and CPGM sequences.

| Tissue | Median [30] (range) | Bias (%) QRAPMASTER (95% CI) | Bias (%) MRF-vFA (95% CI) | Median [29] (range) | Bias (%) QRAPMASTER (95% CI) | Bias (%) MRF-vFA (95% CI) |
|--------|---------------------|-----------------------------|---------------------------|---------------------|-----------------------------|---------------------------|
| Gray matter | T1 (ms)          | 1008 [940,1162] (4.9 [2.9, 6.7]) | -4.4 [-5.5, -3.2]         | 1460 [1165,1615] (-12.4 [-13.9, -10.9]) | -20.8 [-21.7, -19.9]         |
|         | T2 (ms)           | 78 [71,92] (23.1 [21.3, 24.8]) | 2.6 [0.6, 5.1]            | 83 [80,86] (7.2 [5.6, 8.9]) | 5.7 [3.5, 8.5]              |
| White matter | T1 (ms)          | 687 [657,735] (-27.1 [-28.9, -25.6]) | -2.6 [-1.9, 5.4]         | 867 [728,954] (-28.4 [-30.5, -26.21]) | -8.0 [-9.9, -5.9]           |
|         | T2 (ms)           | 75 [69,78] (10.7 [8.8, 12.5]) | -18.7 [-21.3, 16.1]      | 75 [72,78] (2.7 [0.8, 4.5]) | -21.3 [-23.2, -19.5]         |

For QRAPMASTER and MRF-vFA the average bias over all the subjects in percentage to the median literature value and the 95% confidence interval of the calculated bias are reported.

In Supplementary Figs. S18 and S19 the CoVs maps for all the volunteers are shown using QRAPMASTER and MRF-vFA respectively. They present low intra-subject variability in gray and white matter.

The results of the GLM analysis are shown in Supplementary Figs. S2 and S3. Supplementary Fig. S2 shows the absolute value of the mean and the standard deviation of the estimated b parameters over the gray matter, white matter and CSF for QRAPMASTER and MRF-vFA, respectively. In both cases, there was a deviation from the mean for T1 values in gray and white matter (>10%) related to the field strength. T2 values were also influenced by the field strength but less (5%-10%). The nPD didn’t reflect the influence of the field strength. For the T1 and T2 estimated values, there was some variability between subjects (5%), especially for CSF. The different day of acquisition introduced very low variation (<3%) for MRF-vFA and QRAPMASTER. Only ‘Day1’ and ‘Day4’ had higher bias for QRAPMASTER in CSF.

Supplementary Fig. S3 shows the b maps associated with the covariance of the 1.5 T system and 3.0 T system. The GLM was designed with the parameters for the field strength being complementary (Supplementary Fig. S1). We observed longer T1 values on the 3.0 T system than in the 1.5 T system. In the case of T2 values, MRF-vFA showed longer T2 estimated values on the 1.5 T system than on the 3.0 T system, while QRAPMASTER showed the longest T2 estimated values for the 3.0 T system. The maps are quite uniform along all the brain using QRAPMASTER. Using MRF-vFA, we observed that only T1 map is uniform in all the brain. The nPD is more affected by the field strength from the left inner part.

The repeatability in phantom is good for both methods. This allows reliable longitudinal measurements or population group comparisons since to detect changes along the time and/or differences between groups only low variability is needed, regardless of the bias.

The performance of the different techniques on the phantom demonstrated that QRAPMASTER is substantially more accurate since it had less bias. Regarding repeatability, there were some differences in the CoVs obtained with QRAPMASTER and with MRF-vFA depending on field strength and relaxation parameters. Similar to previous MRF (framework on which MRF-vFA is based) studies, MRF-vFA showed more variability for T2 values than for T1 values (in contrast to QRAPMASTER), likely because of its sensitivity to B0 inhomogeneities and intra-voxel incoherence [18,46]. However, on average, we did not observe significant differences in repeatability between QRAPMASTER and MRF-vFA.

In the in-vivo data, no ground truth values are available. Although the phantom does not reflect all in-vivo phenomena, the performance in the phantom can give confidence in the in-vivo accuracy. This was the case for both techniques since the values for T1 and T2 values were in good concordance with other quantitative techniques [7,29,42]. Although there was some bias to the median literature values, a wide spread in quantitative values, especially in T1, has been reported [45]. When looking in detail, the bias in the in-vivo data was similar to the bias in the phantom for the tubes with T1 and T2 values in the range of brain tissue. Quantitative T1 and T2 maps from both techniques had good contrast, allowing us to distinguish brain anatomy. The T2 estimated values for the brain with MRF-vFA compared to QRAPMASTER followed the trend expected from the phantom analysis, with a larger ratio between gray and white matter for MRF-vFA than QRAPMASTER. Consequently, MRF-vFA had higher contrast in the quantitative T2 maps.
Multiple inversion-recovery pulses are played out in multi-slice sequences \( [38] \). In MRF-vFA no B1 correction is applied and this could cause bias in the performance of QRAPMASTER and MRF-vFA in clinical environments.

Another limitation of this study is the use of systems of a single vendor which does not allow conclusions about reproducibility across vendors. The reproducibility and accuracy of the measurements can vary from one vendor to another due to different implementations \([16,23,55]\).

Future works should focus on the acquisition of patient data in order to evaluate the possible clinical impact of these techniques.
Authorship statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication before its appearance in the Magnetic Resonance Imaging Journal.

Acknowledgments

The authors gratefully acknowledge the support from Rolf F. Schulte providing and installing the pulse sequence.

This work was financed by a Research Grant from GE(B-GEHC-05).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mri.2021.09.004.

References

[1] Deoni SCL. Quantitative Relaxometry of the Brain. Top Magn Reson Imaging 2010; 21:101–13. https://doi.org/10.1097/RMR.0b013e31821821e6.
[2] Badve C, Yu A, Dastmalchian S, Rogers M, Ma D, Jiang Y, et al. MR fingerprinting of adult brain tumors: initial experience. Am J Neuroradiol 2017;38:492–9. https://doi.org/10.3174/ajnr.A5035.
[3] Warmjes JBM, Leinhard OD, West J, Lundberg P. Rapid magnetic resonance quantification on the brain: optimization for clinical usage. Magn Reson Med 2008; 60:320–9. https://doi.org/10.1002/mrm.21635.
[4] Sicherbako V, van den Berg CAT, Moonen CTW, Bartels LW. PLANET: an ellipse fitting approach for simultaneous T1 and T2 mapping using phase-cycled balanced steady-state free precession: ellipse fitting approach for T1 and T2 mapping. Magn Reson Med 2018;79:711–22. https://doi.org/10.1002/mrm.26717.
[5] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duex JF, et al. Magnetic resonance fingerprinting. Nature 2013;495:187–92. https://doi.org/10.1038/nature11971.
[6] Gomez FA, Bounincontri G, Molina-Romero M, Sperl JI, Menzel MI, Menze BH. Accelerated parameter mapping with compressed sensing: an alternative to MR fingerprinting. Honolulu, HI, USA. 2017. p. 6.
[7] Deoni SCL, Peters TM, Rutti BK. High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. Magn Reson Med 2005;53: 237–41. https://doi.org/10.1002/mrm.20314.
[8] Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology standards for quantitative imaging biomarkers. Radiology 2015;277: 813–25. https://doi.org/10.1148/radiol.2015142202.
[9] Hagiwara A, Hori M, Cohen-Adad J, Nakazawa M, Suzuki Y, Kasahara A, et al. Linearity, Bias, Intrascanner repeatability, and interscanner reproducibility of quantitative multidynamic multiecho sequence for rapid simultaneous Relaxometry at 3 T: a validation study with a standardized phantom and healthy controls. Invest Radiol 2019;54:39–47. https://doi.org/10.1097/RLI.0000000000000510.
[10] Di Giuliano F, Minosse S, Picchi E, Marfia GA, Da Ros V, Muto M, et al. Comparison between synthetic and conventional magnetic resonance imaging in patients with multiple sclerosis and controls. Magn Reson Mater Phys Biol Med 2020;33:549–57. https://doi.org/10.1007/s11012-019-00997-9.
[11] Betts AM, Leach JL, Jones BV, Zhang B, Serai S. Brain imaging with synthetic MR in children: clinical quality assessment. Neuroradiology 2016;58:1017–26. https://doi.org/10.1007/s00234-016-1729-9.
[12] West H. Clinical validation of synthetic brain MRI in children: Initial experience. 2017. p. 8.
[13] Tanenbaum LN, Tsirouzis AJ, Johnson AN, Naidich TP, DeLano MC, Melhem ER, et al. Synthetic MRI for clinical neuroimaging: results of the magnetic resonance image compilation (MAGIC) prospective, multicenter, multireader trial. AJNR Am J Neuroradiol 2017;38:1103–10. https://doi.org/10.3174/ajnr.A5227.
[14] Hagiwara A, Andica C, Hori M, Aoki S. Synthetic MRI showed increased myelin partial volume in the white matter of a patient with Sturge-Weber syndrome. Neuroradiology 2017;59:1065–6. https://doi.org/10.1007/s00234-017-1908-x.
[15] Andica C, Hagiwara A, Nakazawa M, Tsuruta K, Takano N, Hori M, et al. The advantage of synthetic MRI for the visualization of early white matter change in an infant with sturge-weber syndrome. MRMS 2016;15:347–8. https://doi.org/10.2463/mrms.ci.2015-0164.

Fig. 8. Maps of the coefficient of variation of nPD, T1, and T2 with QRAPMASTER and MRF-vFA on a 1.5 T system and on a 3.0 T system, averaged over all subjects.
Magnetic Resonance Imaging 83 (2021) 196–207

[16] 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 2018;79:2564–75. https://doi.org/10.1002/mrm.26903.

[17] Buonincontri G, Biagi L, Rietico A, Cecchi P, Cosottini M, Gallagher FA, et al. Multi-site repeatability and reproducibility of MR fingerprinting of the healthy brain at 1.5 and 3.0 T. Neuroimage 2019;195:362–72. https://doi.org/10.1016/j.neuroimage.2019.03.047.

[18] Jiang Y, Ma D, Keenan KE, Stupic KF, Gulani V, Griswold MA. Repeatability of magnetic resonance fingerprinting T1 and T2 estimates assessed using the ISMRM/NIST MRI system phantom: repeatability of MR fingerprinting. Magn Reson Med 2017;78:1452–7. https://doi.org/10.1002/mrm.26509.

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

[20] Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in FMRI. NeuroImage 2003;20:1052–63. https://doi.org/10.1016/S1053-8119(03)00435-X.

[21] Weigel M. Extended phase graphs: Dephasing, RF pulses, and echoes - Pure and simple: extended phase graphs. J Magn Reson Imaging 2015;41:266–75. https://doi.org/10.1002/jmri.24619.

[22] Weigel M. Extended phase graphs: Dephasing, RF pulses, and echoes - Pure and simple: extended phase graphs. J Magn Reson Imaging 2015;41:266–75. https://doi.org/10.1002/jmri.24619.

[23] Hagtaiwa A, Warnantjes M, Hori M, Andica C, Nakazawa M, Kumamaru KK, et al. SyMRI of the brain: rapid quantification of relaxation rates and proton density, with synthetic MRI, automatic brain segmentation, and myelin measurement. Invest Radiol 2017;52:647–57. https://doi.org/10.1097/RLI.0000000000000365.

[24] Tamir JI, Uecker M, Chen W, Lai P, Alley MT, Vasanawala SS, et al. Fessler JA, Sutton BP. Nonuniform fast fourier transforms using min-max interpolation. IEEE Trans Signal Process 2003;51:560–74. https://doi.org/10.1109/TSP.2002.807005.

[25] Santyr E. Magnetization transfer effects in multislice MR imaging. Magn Reson Med 1999;32:206–10. https://doi.org/10.1002/mrm.1910320209.

[26] Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are age-related changes in tissue value properties in children: simultaneous quantification of relaxation times and T2* estimates assessed using the ISMRM/NIST MRI system phantom: repeatability of MR fingerprinting. Magn Reson Med 2017;78:1452–7. https://doi.org/10.1002/mrm.26509.

[27] Neill M, Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, et al. Synthetic MRI for clinical neuroimaging: results of the magnetic resonance image compilation (MAGIC) prospective, multicenter, multireader trial. Am J Neuroradiol 2017;38:1103–10. https://doi.org/10.3174/ajnr.A5227.

[28] Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are age-related changes in tissue value properties in children: simultaneous quantification of relaxation times and T2* estimates assessed using the ISMRM/NIST MRI system phantom: repeatability of MR fingerprinting. Magn Reson Med 2017;78:1452–7. https://doi.org/10.1002/mrm.26509.

[29] Santyr E. Magnetization transfer effects in multislice MR imaging. Magn Reson Med 1999;32:206–10. https://doi.org/10.1002/mrm.1910320209.

[30] Santyr E. Magnetization transfer effects in multislice MR imaging. Magn Reson Med 1999;32:206–10. https://doi.org/10.1002/mrm.1910320209.

[31] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.

[32] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.

[33] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.

[34] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.

[35] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.

[36] de Boer RW. Magnetization transfer contrast part 1: MR physics. 1995.