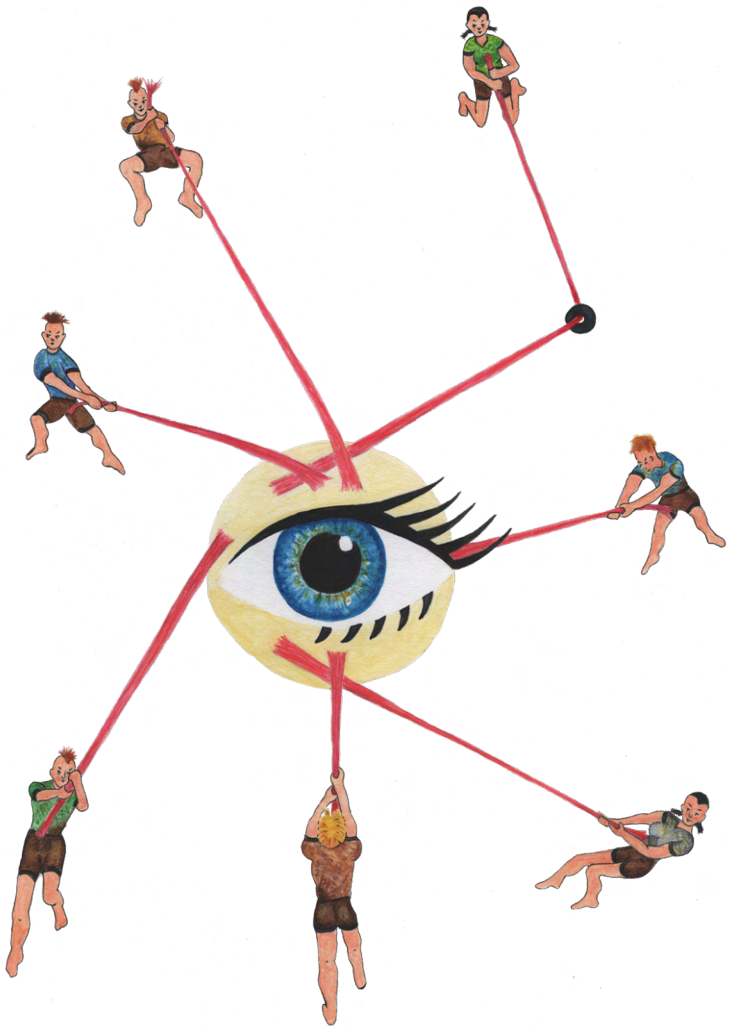
Function and structure of the eye muscles in myasthenia gravis: novel methods to aid in diagnosis and understanding of pathophysiology

Keene, K.R.

1. C.J. Gorter center for high field MRI, Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
2. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands
3. Amsterdam University Medical Center, Amsterdam, the Netherlands
4. Duchenne Center Netherlands, the Netherlands
5. Department of Ophthalmology, Leiden University Medical Center, Leiden, the Netherlands
6. Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, the Netherlands
7. Department of Radiology, University Medical Center Utrecht, Utrecht, the Netherlands



5

T2 relaxation time mapping in healthy and diseased skeletal muscle using extended phase graph algorithms

Magnetic Resonance in Medicine, Nov 2020
doi:10.1002/mrm.28290

Kevin R. Keene[1,2] | Jan-Willem Beenakker[1,5] | Melissa T. Hooijmans[3]
Karin J. Naarding[2,4] | Erik H. Niks[2,4] | Louise A.M Otto[6] | W. Ludo van der Pol[6]
Martijn R. Tannemaat[2] | Hermien E. Kan[1,4] | Martijn Froeling[7]

ABSTRACT

Purpose

Multi-echo spin-echo (MSE) transverse relaxometry mapping using multi-component models is used to study disease activity in neuromuscular disease (NMD) by assessing the T2 of the myocytic component ($T2_{water}$). Current extended phase graph (EPG) algorithms are not optimized for fat fractions above 50% and the effects of inaccuracies in the $T2_{fat}$ calibration remain unexplored. Hence, we aimed to improve the performance of EPG fitting methods over a large range of fat fractions, by including the slice selection flip angle profile, a through-plane chemical shift displacement correction, and optimized calibration of $T2_{fat}$.

Methods

Simulation experiments were used to study the influence of the slice flip angle profile with chemical shift and $T2_{fat}$ estimations. Next, in vivo data from four NMD cohorts were studied for different $T2_{fat}$ calibration methods and $T2_{water}$ estimations.

Results

Excluding slice flip angle profiles or chemical shift displacement resulted in a bias in $T2_{water}$ up to 10 ms. Furthermore, a wrongly calibrated $T2_{fat}$ caused a bias of up to 4 ms in $T2_{water.}$ For the in vivo data, one-component calibration led to a lower $T2_{fat}$ compared to a two-component method, and $T2_{water}$ decreased with increasing fat fractions.

Conclusion

In vivo data showed a decline in $T2_{water}$ for increasing fat fractions, which has important implications for clinical studies, especially in multi-center settings. We recommend using an EPG based model for fitting $T2_{water}$ from MSE sequences with two-component $T2_{fat}$ calibration. Moreover, we recommend including the slice flip angle profile in the model with correction for through-plane chemical shift displacements.

INTRODUCTION

Quantitative MRI shows promising results as a biomarker in the follow-up of disease progression in neuromuscular diseases (NMD).[36] In these diseases, inflammation and progressive fat replacement of muscle tissue are major histological hallmarks. While imaging of the fat fraction using MRI or MR-spectroscopy (MRS) is used to assess disease progression, transverse relaxometry mapping or T2 relaxation time mapping has been applied as an MRI marker for disease activity.[124,196,197]

Quantitative measurements of T2 relaxation time are commonly performed using T2prep sequences[198], multi-acquisition single TE SE sequences, IDEAL-CPMG[199], MR spectroscopy[200] or multi-spin-echo (MSE) sequences[201]. Acquisition times for multi-acquisition sequences are longer[202] (typical acquisition time 10 min) compared to MSE sequences (typical acquisition time: 3 min) and more sensitive to motion artefacts, while MR spectroscopy can only give information for a single muscle. Therefore, MSE is commonly used, as it allows for fast scans with a large field-of-view.[201] To measure the T2 relaxation time from a MSE sequence, a mono-exponential decay can be fitted to the signal intensity as a function of echo time. However, this combined T2 relaxation time, also defined as the 'global T2 relaxation time'[203], represents the combined relaxation of all different components, including the T2 relaxation time of water and fat. In the presence of fat replacement, which is often the case in NMD, this combined T2 is primarily affected by the fat signal, which has a relatively long T2 relaxation time, masking the more subtle changes in T2 due to the other changes in muscle tissue.[41] Therefore, in NMD it is preferred to separate the signal into different components for water ($T2_{water}$) and fat ($T2_{fat}$), where the T2 of the myocytic component ($T2_{water}$) has been proposed as a more accurate biomarker for disease activity.[203]

Assessment of $T2_{water}$ using MSE has been performed using various approaches including fat suppression and modeling. Fat suppression has the advantage that it is available on all scanners, however, it is sensitive to field inhomogeneities and is often unable to suppress the entire fat spectrum[204,205]. As an alternative, different modelling approaches have been proposed to separate the contribution of the water and fat signal. Originally, bi-exponential[206] or tri-exponential[207,208] methods were introduced to separate both contributions to the signal at the successive echo times. However, these models assume a perfect mono-exponential decay for the components, which is generally not observed in vivo. One important deviation from this perfect mono-exponential decay are stimulated echoes that result in a oscillation of the measured signal.[209,210] This oscillation results in an overestimation of the T2, when estimated using an exponential model. As a result, the accuracy and reproducibility of these exponential models are limited.[41]

To overcome these limitations, which are primarily caused by deviations from the assumed Hahn echo sequence due to for example a lower B1, fitting using extended phase graphs (EPG)[211] has been proposed. The EPG concept is a tool for describing the magnetization response in multi-pulse MR sequences, using a Fourier based magnetization description. Similar to exponential models a two component EPG model can be used to account for the fat signal, which has been shown to yield $T2_{water}$ values in the range of the gold standard MRS, and which were largely independent of fat fractions up to 50% in several NMDs.[212,213] However, there are still several aspects that are not taken into account in the currently used EPG analyses.

Firstly, current analyses do not report $T2_{water}$ values above a fat fraction of about 50%[41], while in muscular dystrophies fat fractions above 50% are not an exception[197], thereby limiting the applicability of the method. Secondly, the effect of inaccuracies in the estimation of T2 fat remain unexplored. Determination of the $T2_{fat}$ from MSE sequences is challenging as the spectral components of fat each have different T2 relaxation times which, due to J-coupling, also depends on the echo-spacing.[43,214,215] The current EPG implementations calibrate T2 relaxation time of the fat component using subcutaneous fat[208], thereby assuming that the $T2_{fat}$ in muscle is similar to subcutaneous fat. However, subcutaneous fat contains on average only 90% fat and 10% water[216], rather than 100% fat. Thirdly, the spatial variation of the flip angle across the slice is not taken into account.[41]. Due to the limited duration of the RF-pulses, the excitation and refocusing pulses do not provide a homogeneous flip angle across the excited slice. This is especially the case in T2-mapping using MSE, as short RF-pulses are applied to facilitate short echo spacings.[209,210] Fourthly and finally, the chemical shift displacement between water and fat in the slice direction has to be taken into account. When the slice selection gradient of the excitation pulse is different from the slice selection gradient of the refocusing pulse, the flip angle profiles are not aligned for the protons in fat. Therefore, the fat protons experience a different refocusing than the water protons, resulting in different strength of the stimulated echoes of the fat signal.

In the present work, we aim to improve the performance of EPG fitting methods to determine $T2_{water}$ in muscles of different NMDs with a large range of fat fractions by studying the effect of the incorporation of the flip angle profile with a chemical shift displacement in the slice direction and the assumptions for the calibration of the $T2_{fat}$. We present the importance of including slice flip angle profiles with a chemical shift displacement in the slice direction and correct calibration methods for the T2 of the fat component. Furthermore, we show the performance of the model in four clinical cohorts, which show a gradual decline in $T2_{water}$ for increasing fat fractions.

METHODS

A signal model based on extended phase graphs, including a slice flip angle profile with a chemical shift between water and fat, was used to fit the MSE signal. The fit was performed using a dictionary fitting method as proposed by Marty et al.[41] The performance of this model was studied in simulation experiments. The influence of the slice flip angle profile with the chemical shift and the influence of the assumed $T2_{fat}$ were studied in simulation 1 and 2 respectively. Using in vivo data from four patient cohorts, we performed two additional experiments. In experiment 1 the performance of different calibration methods for the $T2_{fat}$ was examined and in experiment 2 the model was applied to the data, to study the performance and the outcome parameters. All analyses were performed in Matlab (MATLAB 2016a, The Mathworks of Natick, Massachusetts, USA).

EPG-model

The model used for the MSE signal simulation comprises a water and a fat signal simulated with extended phase graphs[211], including the T1, the T2 and the B1. The slice flip angle profiles of the RF-pulses were calculated using Bloch equations at 30 samples along the slice flip angle profile (3x slice thickness).[209,210] Slice angle profiles for a B1 of 70%, 100% and 130% are shown in figure 1A, respectively for the water signal, the fat signal and the combined signal.

Since the slice selection gradients can differ between the excitation and refocusing pulses, the slice flip angle profiles are not aligned for the protons in fat (figure 1). This chemical shift artefact was incorporated in the model for the fat signal, by shifting the fat signal in the slice direction according to:

$$\mathbf{\Delta L} = \frac{\mathbf{1}}{\boldsymbol{\gamma} * \mathbf{G}} * \mathbf{\Delta f} * \mathbf{B0}$$

with ΔL the shift in location in meters, γ the Larmor frequency in hertz/Tesla, G the gradient strength in Tesla/meter, Δf the chemical shift between water and fat in hertz/Tesla and B0 the field-strength of the scanner in Tesla.

Fitting of MSE signal

The measured signal vector $\boldsymbol{S}$ with size $\boldsymbol{N_{echo}}$ (the number of echo's) can be approximated using a bi-component EPG model where the modeled signal vector $\hat{\boldsymbol{S}}$ can be defined as

$$\hat{\boldsymbol{S}} = w\,\boldsymbol{S}_{water} + f\,\boldsymbol{S}_{fat} = [\boldsymbol{S}_{water} \quad \boldsymbol{S}_{fat}] \begin{bmatrix} w \\ f \end{bmatrix}$$

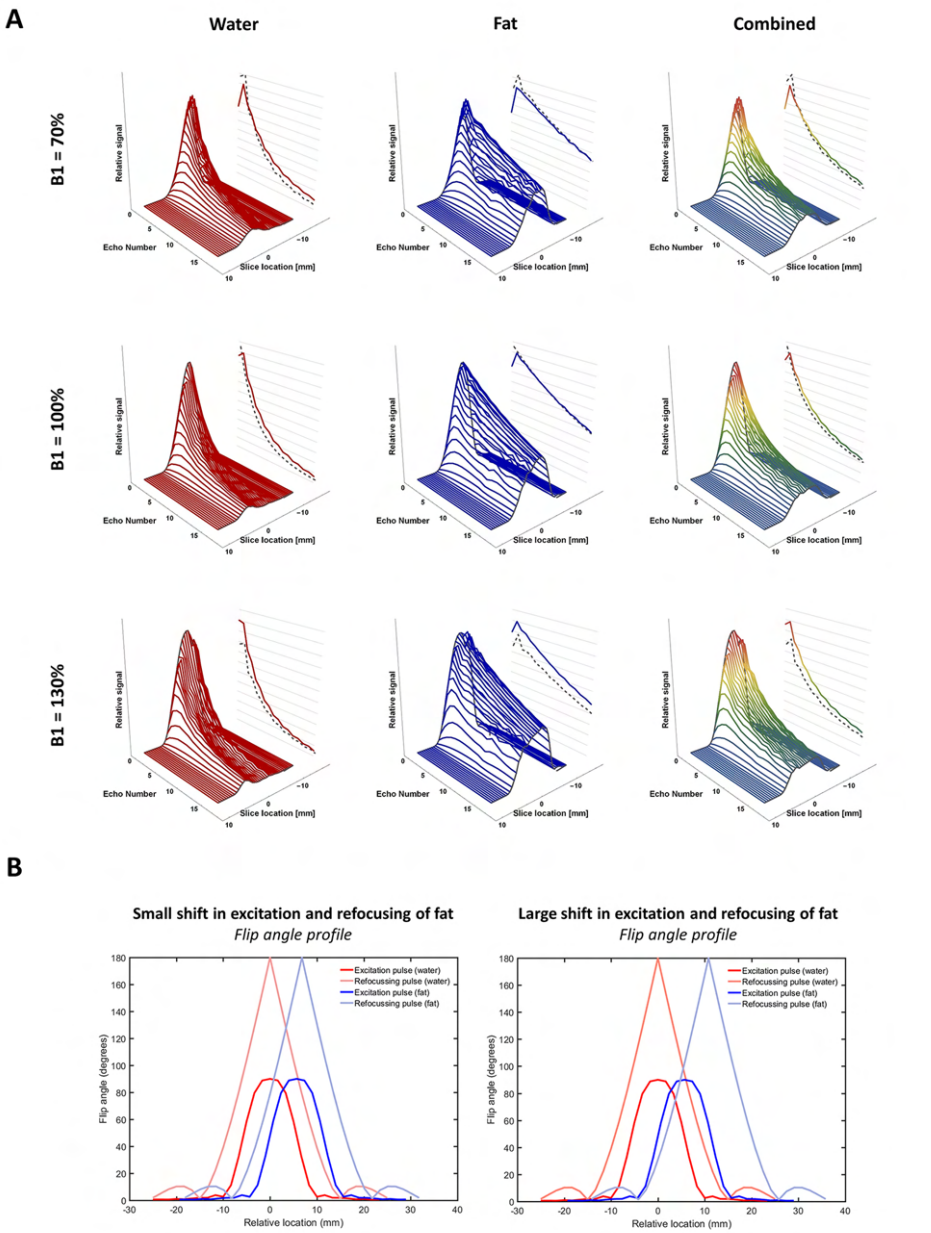


Figure 1: **A:** *Signal profiles for a B1 of 70%, 100% and 130%, from left to right. Noticeable is the asymmetry around a B1 of 100%. The right dotted line in each signal profile represents the signal in the center of the slice, and the right colored line depicts the final signal after averaging all signal contributions.* **B**: *Example flip angle profile of the pulses throughout the slice. On the left an example of a flip angle profile with a minor shift (1.1 mm) between the excitation and refocusing pulse as experienced by the fat is shown and on the right an example of a major shift (5.1 mm).*

With $\boldsymbol{w}$ and $\boldsymbol{f}$ the water and fat signal amplitudes. The signal at each echo of the water and fat component are $\boldsymbol{S}_{water}$ and $\boldsymbol{S}_{fat}$,respectively, and are defined as

$$\boldsymbol{S}_{water} = \sum_{m=1}^{M} EPG\left(T1_{wat}, T2_{wat}, B1, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right)$$

$$\boldsymbol{S}_{fat} = \sum_{m=1}^{M} EPG\left(T1_{fat}, T2_{fat}, B1, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right)$$

With $\boldsymbol{a}_{ex}$ and $\boldsymbol{a}_{ref}$ the flip angle profiles along the slice direction of the excitation and refocusing pulses, respectively, and $\boldsymbol{M}$ the number of samples along the slice profile.
The signal was fitted using a dictionary-based method. The pre-calculated dictionary contains a range of values for $T2_{fat}$, $T2_{water}$ and B1. The T2 parameters ranged between boundaries based on literature values ($T2_{fat}$: 120-200 ms with 2 ms steps and $T2_{water}$: 10 ms-60 ms with 1ms steps) and the B1 value ranged between 50%-140% (with 2% steps), since the incorporation of the flip angle profile breaks the B1 symmetry around 100% for the excitation pulse (see fig 1A).[210] The th dictionary EPG signals $\boldsymbol{d}^{i}_{water}$ and $\boldsymbol{d}^{i}_{fat}$ with $\boldsymbol{T2}^{i}_{wat}$ and $\boldsymbol{B1}^{i}$ are defined as

$$\boldsymbol{d}^{i}_{water} = \sum_{\boldsymbol{m=1}}^{M} EPG\left(T1_{wat}, \boldsymbol{T2}^{i}_{wat}, \boldsymbol{B1}^{i}, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right)$$

$$\boldsymbol{d}^{i}_{fat} = \sum_{m=1}^{M} EPG\left(T1_{fat}, T2_{fat}, \boldsymbol{B1}^{i}, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right)$$

For each dictionary value of $\boldsymbol{T2}^{i}_{wat}$ and $\boldsymbol{B1}^{i}$ the water and fat amplitudes $\boldsymbol{w}^{i}$ and $\boldsymbol{f}^{i}$, and the modelled signal $\hat{\boldsymbol{S}}^{i}$ at the i^{th} dictionary index are estimated using a matrix formulation of the multiple regression model:

$$\begin{bmatrix} \boldsymbol{w}^{i} \\ \boldsymbol{f}^{i} \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}^{i}_{water} \\ \boldsymbol{d}^{i}_{fat} \end{bmatrix}^{-1} \boldsymbol{S}$$

$$\hat{\boldsymbol{S}}^{i} = \begin{bmatrix} \boldsymbol{d}^{i}_{water} & \boldsymbol{d}^{i}_{fat} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}^{i} \\ \boldsymbol{f}^{i} \end{bmatrix}$$

The optimal dictionary values $\boldsymbol{T}^{i}_{2,wat}\boldsymbol{B}^{i}_{1}$ and the corresponding amplitudes of the water and fat signal $\boldsymbol{w}^{i}$ and $\boldsymbol{f}^{i}$ are determined by minimizing the Euclidian norm

$$\operatorname{argmin} \left\| \boldsymbol{S} - \hat{\boldsymbol{S}}^{i} \right\|_{2}$$

To include a water fraction $g(0 \leq g \leq 1)$with $T_{1,watfat}$ and $T_{2,watfat}$ in the dictionary fat signal, d^i_{fat} was defined as

$$\boldsymbol{d}_{fat}^{i} = \sum_{m=1}^{M} (1-g) EPG\left(T1_{fat}, T2_{fat}, \boldsymbol{B1}^{i}, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right) + g\ EPG\left(T1_{watfat}, T2_{watfat}, \boldsymbol{B1}^{i}, \Delta TE, N_{echo}, \boldsymbol{\alpha}_{ex}^{m}, \boldsymbol{\alpha}_{ref}^{m}\right)$$

Simulation experiments

To assess the incorporation of the slice flip angle profile, the chemical shift of the slice profile between water and fat, and the effect of the assumed $T2_{fat}$, the MSE signal evolution was simulated using EPG assuming a known slice flip angle profile (figure 1). In the simulations the $T2_{water}$ was fixed at 30 ms, B1 and fat fractions were randomized between the possible boundaries (B1: 50%-140% and the fat fraction: 0%-100%) and the $T2_{fat}$ was modelled with gaussian distribution (mean of 150 ms and standard deviation 10 ms). The $T1_{water}$ and $T1_{fat}$ were fixed on 1400 ms and 365 ms respectively. Two simulation experiments as described below were performed, one to investigate the effect of slice flip angle profiles, and one to investigate the effect of using wrongly calibrated $T2_{fat}$.

Simulation 1 – The effect of the slice flip angle profile

We conducted two simulations to study the effect of incorporating the flip angle profile and the chemical shift of the slice profile between water and fat signals on the estimated $T2_{water}$. In **simulation 1A,** the data was simulated assuming the full slice flip angle profile and fitting was done with and without assuming the full slice flip angle profile. Next in **simulation 1B,** data was simulated assuming using the full slice flip angle profile with a large and a small chemical shift between water and fat slice profiles. The simulated data was fitted with and without accounting for the chemical shift. The used slice flip angle profiles for the small and large chemical shift are shown in figure 1.

Simulation 2 –The effect of $T2_{fat}$

To study the effect of wrongly calibrated $T2_{fat}$ on the estimated $T2_{water,}$ we conducted a simulation in which a full slice flip angle profile and chemical shift was incorporated. In **simulation 2,** data was simulated with $T2_{fat}$ = 150 ± 10 ms. Next the data was fitted with a correct, an overestimated and an underestimated $T2_{fat}$ of 150 ms, 180 ms and 120 ms, respectively.

In vivo experiments

Study populations and MR scans

To study the effect of T2 calibration methods in vivo, we used data from four different clinical

cohorts that contained MSE data of both patients and healthy controls. The four cohorts contained data from 92 patients and 56 healthy controls. These datasets were acquired in the context of four clinical studies and will be used for methodological purposes in the current work.[217–220] The Medical Ethics Review Committee of the Leiden University Medical Center (LUMC) and the University Medical Center Utrecht (UMCU) approved the studies and all patients or parents provided written informed consent prior to study participation.

All patients were scanned on a 3T Philips MRI system. The first cohort (**cohort 1**) comprised of 18 non-ambulant patients with Duchenne muscular dystrophy (DMD) (17.1±5.1 years, range 9-24 year) and 11 age matched healthy controls (14.7±4.0 years, range 10-25 year), from the LUMC[220]. In this cohort upper-arm scans had been acquired. The second cohort (**cohort 2)** comprised of 22 patients with DMD (9.3±3.1 years, range 5-16 year, 16 ambulant and 6 non-ambulant) and 12 age matched healthy controls (9.7±2.9 years, range 5-14 years), from the LUMC.[219] The third cohort (**cohort 3**) comprised of 23 patients with Becker muscular dystrophy (BMD) (42.7±13.6 years, range 18-67 year, all ambulant) and 13 age matched healthy controls (43.0±13.7 years, range 21-63 year), also from the LUMC[217]. In both cohort 2 and cohort 3 lower-leg scans had been acquired. The fourth and final cohort (**cohort 4**) comprised of 29 patients with spinal muscular atrophy (SMA) (29.6±18.0 years, range 17-71 years, 48% male, 14 patients had SMA type 2 and 15 patients had SMA type 3) and 20 age matched healthy controls (37.9±12.8 years, range 17-71 years, 40% male), from the UMCU[218]. In this cohort upper-leg scans had been acquired. In each data set, regions of interest were manually drawn on 5 slices that consisted of the entire cross-sectional muscle compartment without the bone and subcutaneous fat (figure 2). The sequence parameters for each cohort are depicted in table 1.

With the data of these cohorts we performed two experiments as described below. The first experiment aimed to compare three different $T2_{fat}$ calibration methods using voxels from the subcutaneous fat. The second experiment evaluated the effect of these fat calibration methods on the $T2_{water}$ estimation.

Calibration of $T2_{fat}$

The $T2_{fat}$ was calibrated on subcutaneous fat using a fat mask based on the last echo in the echo sequence.[208] The mask was determined by thresholding the last echo in the MSE sequence with the mean signal intensity (figure 2). The signal was fitted for each voxel using the library method described above, fixing the $T2_{water}$ and the water/fat quantities as described below. The $T2_{fat}$ was obtained by averaging all fitted voxel in the fat mask.

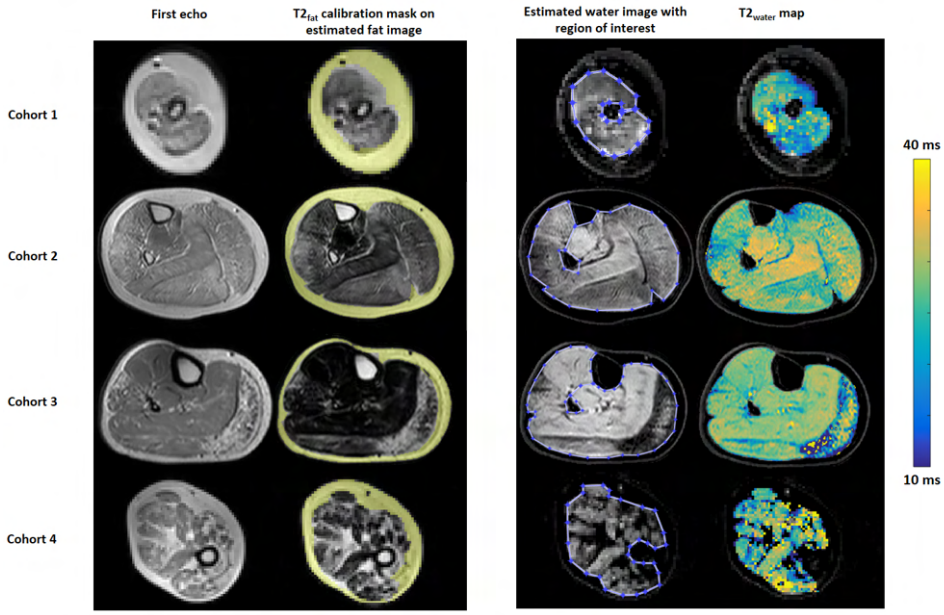


Figure 2. *Example raw data and maps from one patient in all four cohorts. From left to right the signal for the first echo, the estimated fat map with the $T2_{fat}$ calibration mask, the estimated water map with the region of interest and the $T2_{water}$ map are depicted.*

Experiment 1 – in vivo calibration of $T2_{fat}$

The $T2_{fat}$ of the subcutaneous fat was estimated using three different methods. The first, **method A,** assumed one fat component with a 100% fat fraction. The second, **method B**, estimated $T2_{fat}$ using a fixed fat fraction of 90% (based on literature values[221]) and a water component of 10%, where both relaxation times are fitted in the model. In the third and last method, **method C**, calibration was done using a fixed fat fraction of 90% and a fixed $T2_{water}$ in the fat of 20 ms, to stabilize the fit with low water signal.

Experiment 2 – In vivo T2 estimation

To study the performance of the different methods of calibration, we compared the outcome parameter $T2_{water}$ for 5 slices of the patient cohorts using **method A, B and C** as described above. In summary, with **method A** the $T2_{fat}$ is calibrated assuming one fat component (fat fraction 100%), with **method B** the calibration is performed using a fixed fat fraction of 90% and a 10% water-fraction with a fitted $T2_{water}$ and with **method C** the calibration is done using a fixed fat fraction of 90% and a 10% water-fraction with $T2_{water}$ of 20 ms. In the last two methods fitting is performed with the calibrated $T2_{fat}$ that also includes a 10% water

Table 1. *Acquisition and pulse parameters for the four cohorts. Cohort 1 included Duchenne's muscular dystrophy (DMD) patients with healthy controls (HC), cohort 2 included DMD patients with healthy controls, cohort 3 included Becker muscular dystrophy (BMD) patients with healthy controls and cohort 4 included spinal muscular atrophy (SMA) patients and healthy controls.*

	Cohort 1 ***DMD: n=18*** ***HC: n=11***	**Cohort 2** ***DMD: n=22*** ***HC: n=12***	**Cohort 3** ***BMD: n=23*** ***HC: n=13***	**Cohort 4** ***SMA: n=29*** ***HC: n=20***
Anatomical location	*Upper arm*	*Lower leg*	*Lower leg*	*Upper leg*
Sequence parameters				
First echo time	*8.0 ms*	*8.0 ms*	*8.0 ms*	*7.6 ms*
Time between echoes	*8.0 ms*	*8.0 ms*	*8.0 ms*	*7.6 ms*
Number of echoes	*17*	*17*	*17*	*17*
Repetition time	*3000 ms*	*3000 ms*	*3000 ms*	*4598 ms*
Slice thickness	*10 mm*	*10 mm*	*10 mm*	*6 mm*
Slice gap	*20 mm*	*20 mm*	*20 mm*	*0 mm*
Slices	*5*	*5*	*5*	*13*
Resolution	*2.0 mm x 2.0 mm*	*1.6 mm x 1.6 mm*	*1.6 mm x 1.6 mm*	*3.0 mm x 3.0 mm*
*ΔChemical shift**	*0.1 mm*	*1.1 mm*	*1.1 mm*	*0.9 mm*
Excitation pulse				
Bandwidth	*1000 Hz*	*1000 Hz*	*1000 Hz*	*768 Hz*
Gradient strength	*1.78 mT/m*	*1.78 mT/m*	*1.78 mT/m*	*2.28 mT/m*
Refocusing pulse				
Bandwidth	*632 Hz*	*632 Hz*	*632 Hz*	*486 Hz*
Gradient strength	*1.80 mT/m*	*1.48 mT/m*	*1.48 mT/m*	*1.90 mT/m*

** Chemical shift displacement difference in slice direction between excitation and refocusing pulse for fat.*

component with respectively the calibrated $T2_{water}$ and a $T2_{water}$ of 20 ms. Example $T2_{water}$ maps are shown in figure 2.

Statistics

For the simulation experiments, the relation between simulated fat fractions and fitted $T2_{water}$ values and fitted fat fractions was obtained using LOWESS-regression. For analysis a linear regression analysis was performed, and the slope and intercept were reported. The voxels that were fitted with boundary values ($T2_{water}$ shorter than 10 ms and longer than 60ms) were excluded from statistical analysis and depicted in red in the figures. For the calibration experiment (**experiment 1**), smoothed histograms were generated for all calibrated values. Average $T2_{water}$ values between different calibration methods were compared using paired T-tests. $T2_{fat}$ values within the different calibrations were compared using a Kruskal-Wallis test with post-hoc comparison using Bonferroni correction. For the in vivo fitting experiment (**experiment 2**) LOWESS regression was used to assess the relation between $T2_{water}$ and fat fractions.

Data availability

The used Matlab scripts and one example dataset presented in this article will be made available at the request of a qualified investigator. Requests should be made to K. R. Keene (k.r.keene@lumc.nl). A Mathematica version of the fitting algorithm is available at Github (https://github.com/mfroeling/QMRITools).

RESULTS

Simulation 1: The effect of the slice flip angle profile

In **simulation 1A,** data was simulated assuming the slice flip angle profile and fitted with and without taking the flip angle profile into account. The results of the fit are depicted in figure 3. Without incorporation of the flip angle profile, the results show an overestimation of $T2_{water}$, with a median of 40 ms (5^{th}-95^{th} percentile: 34-57 ms), and with many voxels on the dictionary boundaries (30%). The results also show an underestimation of the fitted fat fraction for the entire range of values (intercept: -2.26%, slope: 0.87). However, fitting with incorporation of the flip angle profile resulted in a median $T2_{water}$ of 30 ms (5^{th}-95^{th} percentile: 26 - 35 ms), and only 1.5% of fits were on dictionary boundaries. The correlation between the simulated and fitted fat fraction show an intercept of 0.3% and a slope of 0.99.

In **simulation 1B,** data was simulated assuming a small and a large chemical shift between water and fat, and fitting with and without assuming the chemical shift (figure 2B). The results for the fits simulated and fitted with the small and large chemical shift are comparable ($T2_{water}$: median: 30 ms [5^{th}-95^{th} percentile: 26 – 35 ms], on dictionary boundaries: 1.5%; fat fraction: intercept: 0.3%, slope: 0.99), and show no underestimation or overestimation of the $T2_{water}$ and the fat fraction (figure 4). However, when the chemical shift is not included in the analysis, $T2_{water}$ was overestimated with a median of 32ms (5^{th}-95^{th} percentile 28 – 51 ms) and 34ms (5^{th}-95^{th} percentile: 30 – 53 ms) for the small and large chemical shift respectively. The fat fraction is globally underestimated, and the fit is increasingly unstable for higher fat fractions (small shift: intercept: 0.72%, slope: 0.92; high shift: intercept: -1.28%, slope: 0.79).

Simulation experiment 2 – The influence of the T2fat

In **simulation 2A,** data was simulated with a known $T2_{fat}$ and fitted with an underestimated, correct and overestimated $T2_{fat}$ (figure 5). A baseline was established by fitting the values with an average $T2_{fat}$, with the median of the $T2_{water}$ at 30 ms (5^{th}-95^{th} percentile: 23 – 38 ms) with 4% of voxels on the dictionary boundaries and no apparent error in the fitted fat fraction (intercept: 0.6%, slope: 1.0). With an underestimation of the $T2_{fat}$, the $T2_{water}$ is underestimated and the fat fraction is overestimated, and the fit is more unstable ($T2_{water}$: median: 26ms (5^{th}-

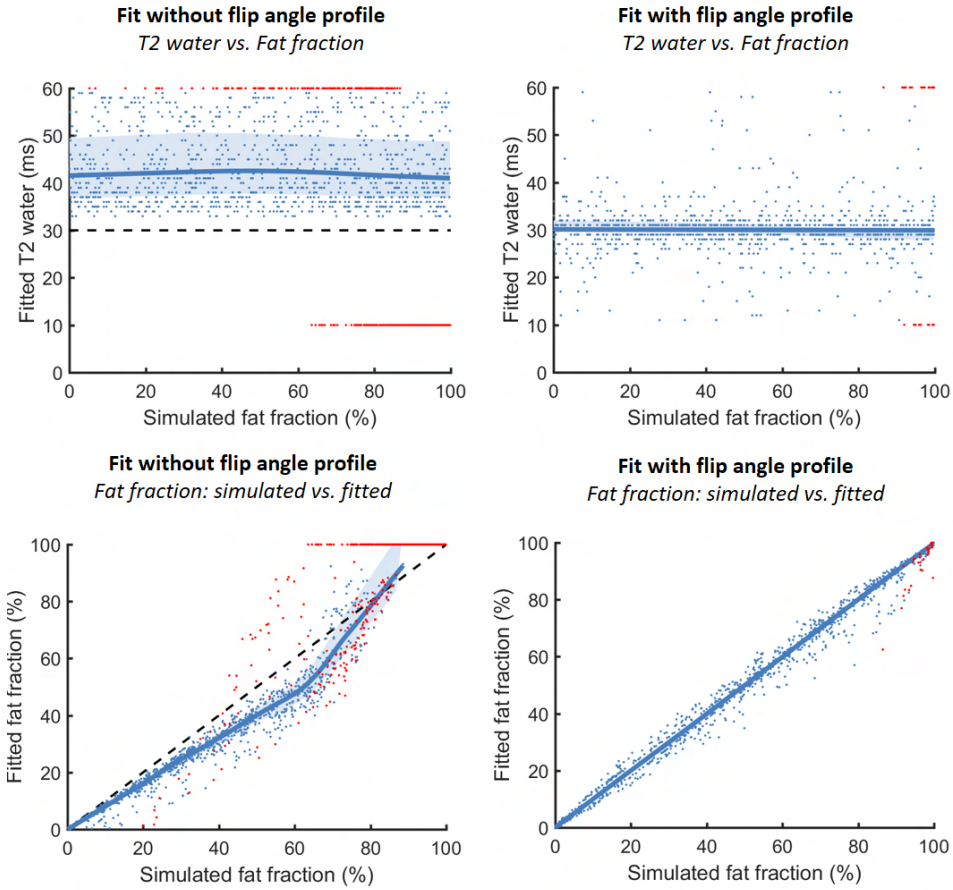


Figure 3. **Simulation 1A** *shows the relationship between the simulated FF and the fitted $T2_{water}$ for simulation without incorporation of the flip angle profile ($T2_{water}$: median: 40 ms [34 ms – 57 ms], on dictionary boundaries:30%) and fitted with incorporating the flip angle profile (top row) ($T2_{water}$: median: 30ms [26 ms – 35 ms], on dictionary boundaries: 1.5%). In the bottom row, the relation between the simulated fat fraction and fitted fat fraction is shown without and with flip angle profile. The values that fitted on dictionary boundaries are depicted in red and the correctly fitted values are depicted in blue. A reference line is shown at a $T2_{water}$ of 30 ms and a unity line is shown for the fat fractions.*

95th percentile: 14 – 30 ms), on dictionary boundaries: 26%; fat fraction: intercept: 2.71%, slope: 1.11). With an overestimation of the $T2_{fat}$, the $T2_{water}$ is overestimated and the fat fraction is underestimated ($T2_{water}$: median: 33 ms (5th-95th percentile: 30 – 51ms), on dictionary boundaries: 9%; fat fraction: intercept: 0.32%, slope: 0.84).

Experiment 1 – In vivo calibration of $T2_{fat}$

The $T2_{fat}$ values for the three described methods to calibrate the $T2_{fat}$ are depicted as smoothed histograms in figure 6, separated for all four clinical cohorts. In **method A,** where the calibration was done assuming one fat component (fat fraction 100%), a median $T2_{fat}$ of 137 ms, 132 ms, 131 ms, and 144 ms was found for cohort 1 to 4 respectively. The $T2_{fat}$ differed significantly between cohorts ($p<0.001$) except between cohort 2 and 3, which were acquired with the same sequence. Using **Method B,** the calibration was done using a fixed fat fraction of 90% and a fixed water component of 10%, of which both the relaxation times were fitted. This method results in a median $T2_{fat}$ of 149 ms, 141 ms, 140 ms and 158 ms, which again differed significantly ($p<0.001$) except between cohort 2 and 3. The water component had a median $T2_{water}$ around 27 ms, 24 ms, 24 ms and 22 ms for cohorts 1 to 4. The histogram for cohort 1 shows that a portion of the values are over 30 ms, possibly representing the $T2_{water}$ of muscle. In the last experiment (**method C**) the calibration was performed using a fixed fat fraction of 90% and a $T2_{water}$ component in the fat of 20ms. This method results in a median $T2_{fat}$ of 150 ms, 142 ms, 140 ms and 159 ms for cohorts 1 to 4 respectively, which is similar to the results of experiment 1B (also statistically different between all cohorts except cohort 2 and 3, $p<0.001$).

The one-component fitting model leads to a lower $T2_{fat}$ compared to the two-component fitting model. Notable is the difference of approximately 10ms in median $T2_{fat}$ between the cohorts for all methods, except for cohort 2 and 3 which have similar $T2_{fat}$ values. Comparing method B and C seem to result in similar $T2_{fat}$ values.

For cohort 1 there was a difference in $T2_{fat}$ for the two component calibrations between the healthy controls and patients, explaining the difference in the histograms with method B and C. In this cohort the average calibrated $T2_{fat}$ was 139 ms for healthy controls and 155 ms for the patients in cohort 1 (DMD arm scans), while for the other cohorts the histograms of the healthy controls and patient overlap, for example for cohort 3 (BMD leg scans) the average calibrated $T2_{fat}$ was 140 ms for both the healthy controls and BMD patients.

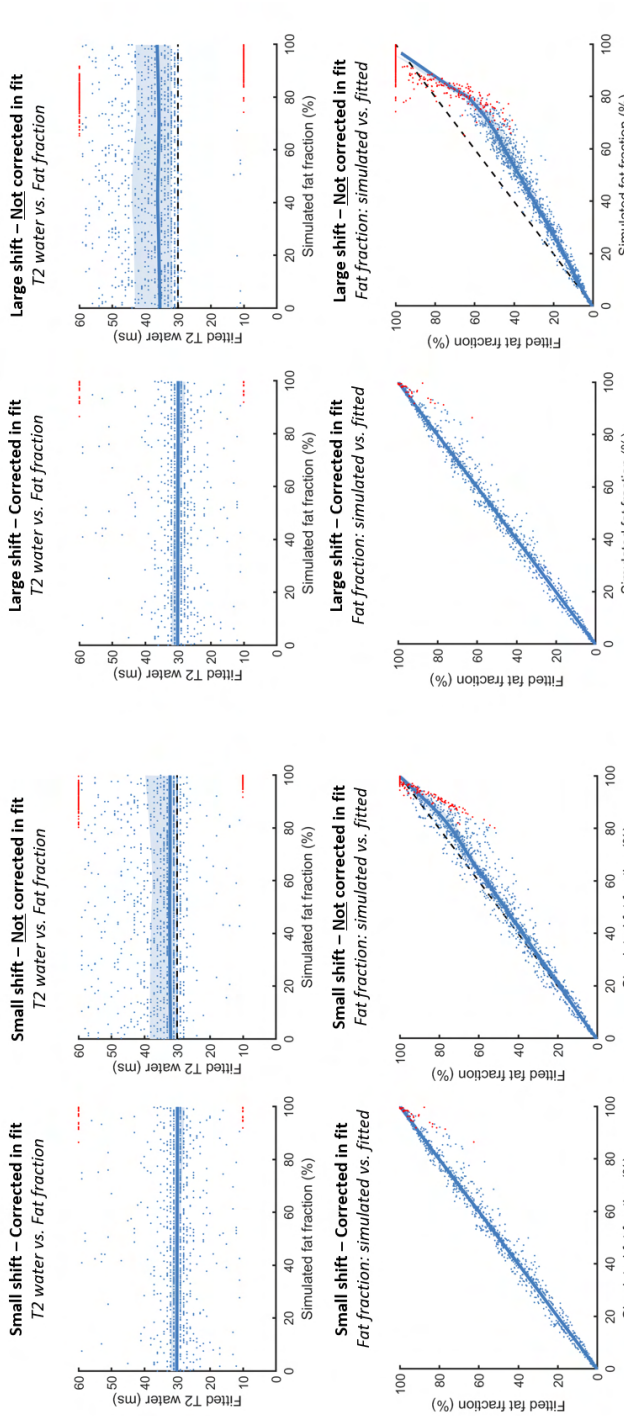


***Figure 4.* Simulation 2** *showing the effect of simulating data with a chemical shift in the location of the pulses and fitting it with or without incorporating the chemical shift of the slice. The four figures on top show the simulation with a small shift ($T2_{water}$: median: 32 ms (28 ms – 51 ms), on dictionary boundaries: 6.8%; fat fraction: intercept: 0.72%, slope: 0.92) and the four figures on bottom show the simulation with a large shift ($T2_{water}$: median: 34 ms (30 ms – 53 ms), on dictionary boundaries: 22%; fat fraction: intercept: -1.28%, slope: 0.79). The values that fitted on dictionary boundaries are depicted in red and the correctly fitted values are depicted in blue. A reference line is shown at a $T2_{water}$ of 30 ms and a unity line is shown for the fat fractions.*

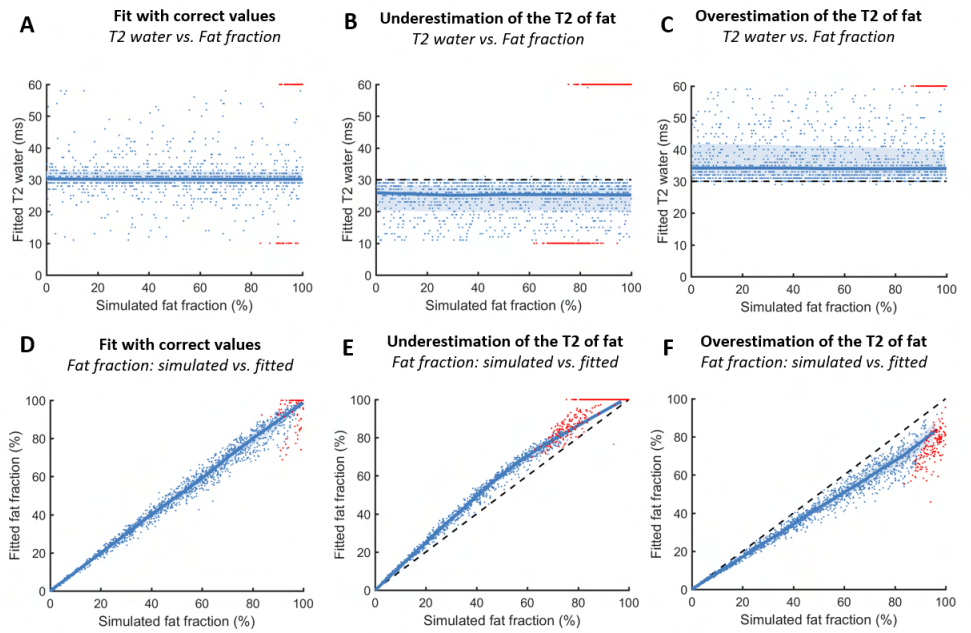


Figure 5. **Simulation 2** *shows the effect of fitting simulated data with different assumed $T2_{fat}$ values. The top row shows the $T2_{water}$ for assuming the correct $T2_{fat}$ ($T2_{water}$: median: 30 ms [23 ms – 38 ms], on dictionary boundaries: 4%, fig. A), an underestimation of $T2_{fat}$ ($T2_{water}$: median: 26 ms [14 ms – 30 ms], on dictionary boundaries: 26%, fig. B) and an overestimation of $T2_{fat}$ ($T2_{water}$: median: 33 ms [30 ms – 51 ms], on dictionary boundaries: 9%, fig C.). Fig. D, E and F show the relationship between the simulated FF and fitted FF. The values that fitted on dictionary boundaries are depicted in red and the correctly fitted values are depicted in blue. A reference line is shown at a $T2_{water}$ of 30 ms and a unity line is shown for the fat fractions.*

Table 2. *The average $T2_{water}$ and fat fractions (+ standard deviations) for the four cohorts divided in healthy controls and patients. The results are shown for the one-component calibration (method A) and the two-component calibration (method C). Regions of interest included the entire muscle compartment without the bone and subcutaneous fat.*

		One-component calibration		Two-component calibration	
Cohort		**Fat fraction**	**$T2_{water}$**	**Fat fraction**	**$T2_{water}$**
1. DMD arm	*Healthy controls*	*4.7 ± 1.7%*	*27.4 ± 1.0 ms*	*4.9 ± 1.7%*	*27.5 ± 1.0 ms*
	DMD patients	*55.8 ± 15.0%*	*19.6 ± 4.7 ms*	*62.0 ± 10.4%*	*24.0 ± 2.2 ms*
2. DMD leg	*Healthy controls*	*9.2 ± 3.7%*	*28.8 ± 0.6 ms*	*9.9 ± 3.0%*	*29.2 ± 0.6 ms*
	DMD patients	*39.0 ± 22.6%*	*25.5 ± 4.9 ms*	*41.4 ± 24.2%*	*28.8 ± 3.9 ms*
3. BMD	*Healthy controls*	*11.9 ± 2.5%*	*28.8 ± 0.7 ms*	*14.7 ± 9.1%*	*29.2 ± 0.7 ms*
	BMD patients	*29.1 ± 17.2%*	*27.0 ± 2.9 ms*	*30.1 ± 18.5%*	*27.8 ± 2.2 ms*
4. SMA	*Healthy controls*	*17.3 ± 7.2 %*	*28.5 ± 0.7 ms*	*18.1 ± 6.9%*	*29.2 ± 0.7 ms*
	SMA patients	*84.1 ± 19.4%*	*20.6 ± 4.2 ms*	*86.3 ± 24.7%*	*27.1 ± 3.4 ms*

Experiment 2 – In vivo T2 estimation

The $T2_{water}$ values differed less than 1ms between method B and C for all patients (mean difference 0.7ms, $p<0.001$, paired t-test). Method B resulted in unphysiologically high $T2_{water}$ values of the fat component (above 35ms) in healthy controls, showing that this calibration method is less stable. Therefore method C was used to compare one-component versus two-component calibrations. The average $T2_{water}$ for healthy controls was comparable between these two methods of calibration. The $T2_{water}$ was 27.4 ms, 28.8 ms, 28.7 ms and 28.5 ms for the one-component calibration and 27.5 ms, 29.2 ms, 29.2 ms and 29.2 ms for the two-component calibration for cohort 1 until 4 respectively. For the patients, the $T2_{water}$ in the one-component calibration is lower ($p<0.001$, paired t-test) than the two-component calibration. The $T2_{water}$ was 19.6 ms, 25.5 ms, 26.6 ms and 20.6 ms for the one component calibration, and 24.0 ms, 28.8 ms, 27.7 ms and 27.1 ms for the two-component calibration, for cohort 1 to 4. The fat fractions were higher in patients than in healthy controls (table 2).

In figure 7 the relation between fat fraction and $T2_{water}$ is shown for each cohort and the three calibration methods. Using the one-component calibration method, with increasing fat fraction the $T2_{water}$ decreases for all cohorts. With the two-component calibrations, the negative correlation between $T2_{water}$ and fat fractions is reduced in cohorts 1 to 3 (DMD and BMD cohorts), especially in the fat fraction range < 50%, and absent in cohort 4 (SMA cohort).

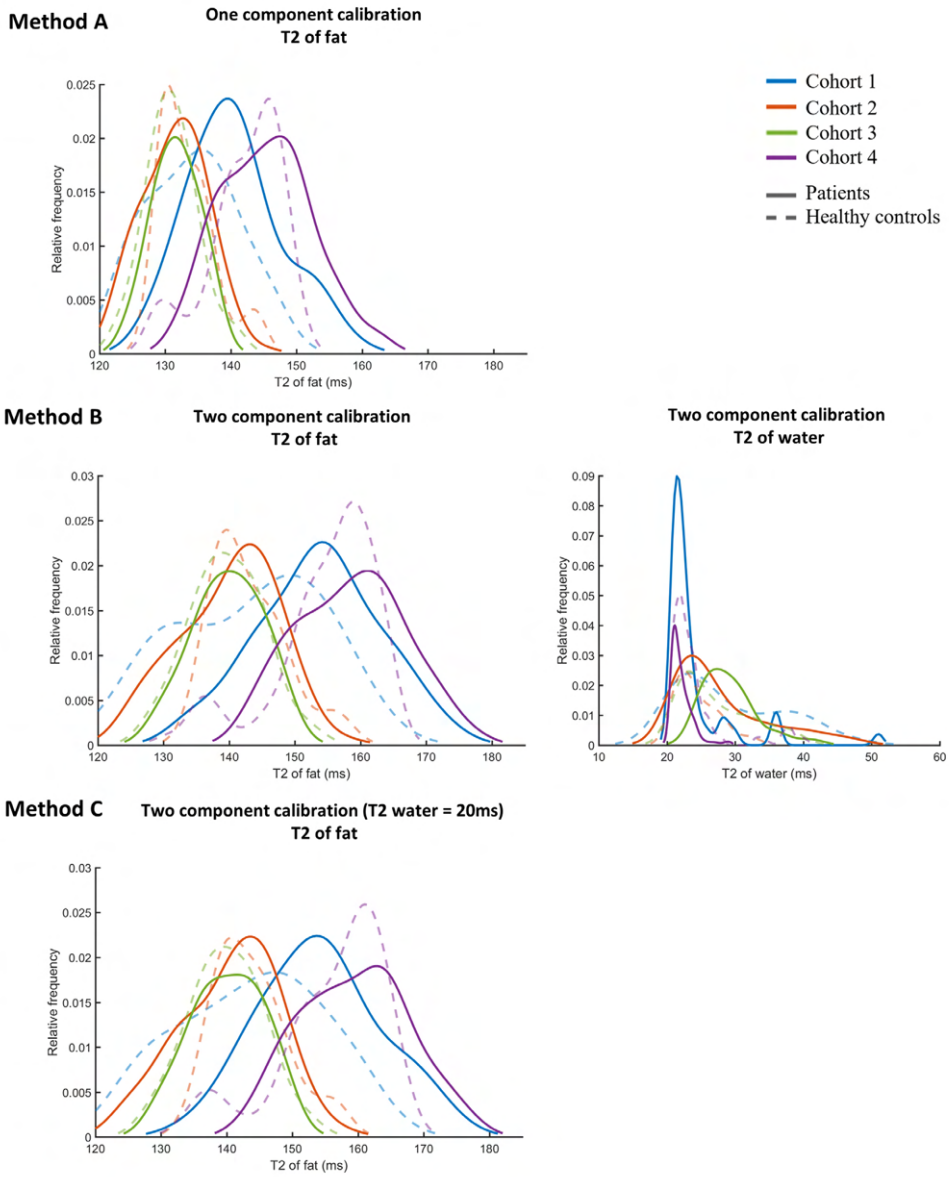


Figure 6. *Outcome of the calibration of the $T2_{fat}$ on subcutaneous fat for the methods A, B and C. The values for the patients in the cohorts are shown as a solid line and the values for the healthy controls are shown as a dashed line. Method A with one component calibration leads to a lower $T2_{fat}$ (137 ms, 132 ms, 131 ms, 144 ms, for cohort 1-4 respectively) than method B with two component calibration (149 ms, 141 ms, 140 ms, 158 ms, for cohort 1-4 respectively). The calibrated $T2_{fat}$ is comparable for method B without a fixed $T2_{water}$ and method C with a fixed $T2_{water}$ (150 ms, 142 ms, 140 ms, 159 ms, for cohort 1-4 respectively).*

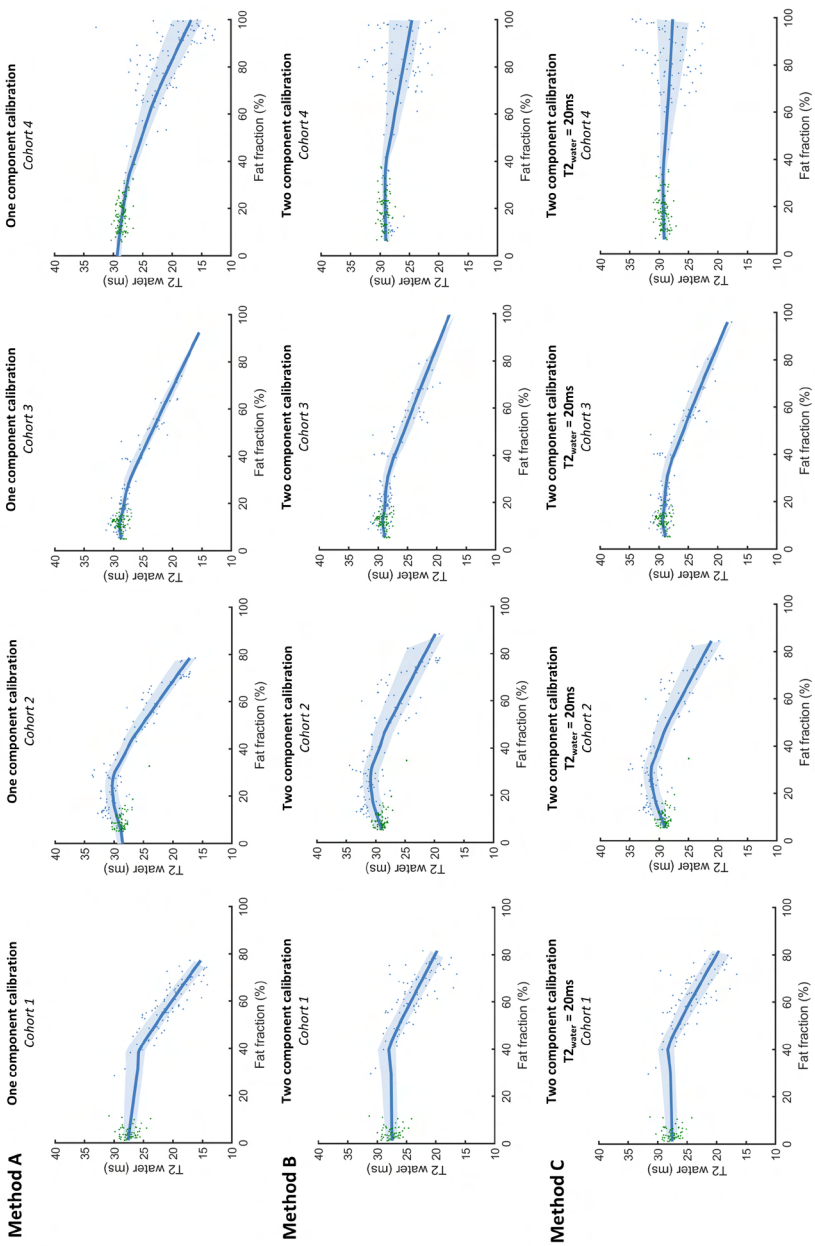


Figure 7. *The association between the fat fraction and the T2 of the water component in vivo in four cohorts (cohort 1-4 from left to right) for the two calibration methods A (top row), B (middle row) and C (bottom row). Healthy controls are depicted in green and patients in blue. Values are shown for each 5 slices per individual.*

DISCUSSION

Multi-echo spin-echo transverse relaxometry mapping using multi-component models are used to study disease activity in neuromuscular disease. Recently, an EPG model has been introduced to obtain separate T2 values for water and fat, accounting for B1 and stimulated echoes.[41] We improved this model and showed the importance of including slice flip angle profiles with a chemical shift displacement in the slice direction. Different calibration methods for the T2 of the fat component showed that the assumption of the T2 of the fat component has a large influence of the estimation of $T2_{water}$. Finally, we studied the performance of the model in four clinical cohorts, which showed a gradual decline in $T2_{water}$ for increasing fat fractions.

Our simulations showed that not including the flip angle slice profile introduced an overestimation of the $T2_{water}$ up to 10 ms and an underestimation of the fat fraction up to 20%. Accordingly, without accounting for the flip angle slice profile the calibrated $T2_{fat}$ is also overestimated. Including the flip angle profile also improved the stability of the fit in higher fat fractions, with fewer voxels reaching the dictionary boundaries. With better defined slice profiles, which can be obtained by increasing RF duration, this effect can likely be reduced. Additionally, simulations showed the importance of including the slice profile chemical shift in EPG simulations for the fat signal, when the gradient strength of the slice selection gradient differs for the excitation and refocusing pulse. The water fat shift in the slice-encoding direction is larger at high field, with thin slices and with large differences between the slice encoding gradients of the excitation and refocusing RF pulses. Including the flip angle profile and its chemical shift makes $T2_{water}$ values comparable between different sequences and different MRI systems, improving comparison between cohorts in literature and in multi-center trials. These simulations show that it is essential to be aware of the sequence used in clinical studies, especially in the context of multicenter studies where exact sequence parameters between MR systems and/or vendors are likely different.

The assumption of the T2 of the fat component has a large influence of the estimation of $T2_{water}$. Assuming one general $T2_{fat}$ without specific calibration is not recommended, since the T2 of the fat component can differ for different scanners and sequences due to J-coupling.[43] The $T2_{fat}$ can automatically be calibrated on the subcutaneous fat as a reference for fatty tissue. From in vivo studies[221] and Dixon scans of subcutaneous fat, we know that the fat fraction in subcutaneous fat is around 90%. Our simulations show that the $T2_{fat}$ is underestimated by approximately 10ms when using a one component model for fat calibration. Therefore, performing a two-component fit improves the accuracy of the estimated $T2_{fat}$ and therefore the accuracy of the estimated $T2_{water}$ and fat fraction.

In the in vivo data we studied, the calibrated $T2_{fat}$ differed between sequences, which could be explained by differences in scan and sequence parameters. The echo time, slice thickness and profiles all influence J-coupling of fat protons which might result in different apparent $T2_{fat}$ values. Additionally, between subjects within the same cohort there was a variation in the $T2_{fat}$, possibly reflecting differences in composition of subcutaneous fat between subjects and MRI acquisition parameters like B1. For cohort 1 with upper-arm scans of DMD patients and healthy controls, there was a difference in the calibrated $T2_{fat}$ between healthy controls and patients (figure 6). This can probably be attributed to partial volume effects, as the thickness of the subcutaneous fat is covered by a relatively sparse number of voxels in the arms due to the low resolution of these scans, and the subcutaneous fat in the arm only covered two to four voxels in most healthy controls. To allow accurate and robust calibration of $T2_{fat}$ it is essential that there are enough pure subcutaneous fat voxels, preferably equally distributed over slices and in regions with varying B1. As such, basing the calibration on a small manual region is not recommended.

The in vivo analyses of four cohorts, with healthy controls and patients with different neuromuscular diseases, was performed with two proposed calibration methods. Assuming the subcutaneous fat to be one component of only fat, we saw a gradual decline in the $T2_{water}$ with increasing fat fractions for all cohorts. Assuming subcutaneous fat to consist of 90% fat and 10% water, the decline in $T2_{water}$ for higher fat fractions decreased for DMD and BMD, and was even absent in the SMA cohort (cohort 4). The decrease of $T2_{water}$ with increase of fat fraction in our cohorts is consistent with the $T2_{water}$ measured with the gold standard spectroscopy in patients due to other neuromuscular disease with fatty replacement.[222] Spectroscopy based $T2_{water}$ measurements are typically performed using voxel localization containing both muscle and intramuscular fat. Therefore, similar to the model here, spectroscopy based $T2_{water}$ estimations are likely biased towards lower $T2_{water}$ as well, with increasing contributions of the water component of fat, which has a lower $T2_{water}$ compared to muscle.

While the scope of this work was methodological optimization of multi-component MSE fitting, data from the two-component fitting method can be used to look at clinical differences (figure 7; method C). The dependence of $T2_{water}$ on the fat fraction was different in the SMA cohort compared to the DMD and BMD cohorts. This may be explained by differences in pathophysiology between SMA and dystrophinopathies (DMD/BMD). SMA is a disease characterized by muscle atrophy, presumably caused by the effects of denervation secondary to motor neuron degeneration, in addition to some fat replacement of muscle tissue.[223] In BMD and DMD pathophysiology is primarily characterized by replacement of muscle tissue by fat and endomysial fibrosis, of which the latter can reach up to 35% of muscle biopsy

areas in histological studies.[224] The $T2_{water}$ of fibrotic tissue is below 10 ms[39,225], possibly explaining the $T2_{water}$ decrease in severely affected DMD and BMD individuals with high amounts of fat replacement. Alternatively, susceptibility differences between muscle water and fat and have also been mentioned as explanations.[222] However, to fully explain this difference further work is needed.

There are some limitations of the used model. We assumed that the fat in tissue contributes to the signal as one mono-exponential component. However, fat contains several components with different relaxation times and J-coupling interactions.[43] These J-coupling effects influence the signal in a complex matter by stimulated echoes, possibly causing the relaxation to deviate from pure mono-exponential behavior. Furthermore, we assume a fixed water fraction in the fat calibration, and for fitting stability for this fraction we fixed the T2 relaxation of this water component to 20 ms. The assumption of a fixed $T2_{water}$ for the water component in fat resulted in the most stable fit for all investigated cohorts, even though the calibration method with a two-component calibration in fat shows a 0.7ms difference in $T2_{water}$. In addition, in our final recommended model (C) the assumptions of ~10% fat contribution and a water T2 in fat of ~20 ms are supported by previous work.[221,226] In future work in different cohorts, the $T2_{water}$ of the fat component could be included in the fit and the stability of that model could be studied. Additionally, for the calibration of $T2_{fat}$, we assume that the fatty replacement in muscle tissue has the same composition and relaxation parameters as subcutaneous fat. However, differences exist in the composition of different fat tissues.[216] More research is needed to quantify the composition of fat that infiltrates the muscle in neuromuscular diseases to mitigate potential biases due to these assumptions.

To conclude, we recommend using an EPG based model for fitting $T2_{water}$ from the MSE signal with calibration of $T2_{fat}$ assuming two components. Moreover, we recommend including the slice flip angle profile in the model which includes chemical shift displacements. In vivo data showed a gradual decline in $T2_{water}$ for increasing fat fractions, with important implications for clinical studies using $T2_{water}$ as an outcome parameter. Using these recommendations, $T2_{water}$ measurements will be more reliable and will allow for better comparison of values between centers and diseases.