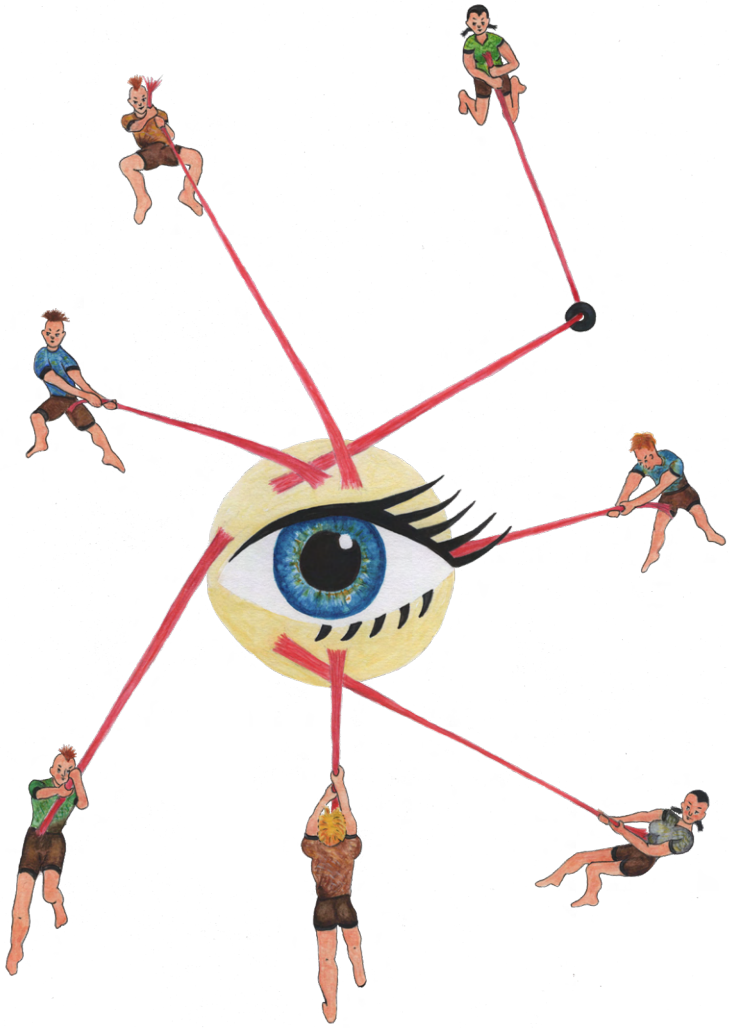
Function and structure of the eye muscles in myasthenia gravis: novel methods to aid in diagnosis and understanding of pathophysiology

Keene, K.R.

1. CJ Gorter center for high field MRI, Department of Radiology, Leiden University Medical Center, Leiden, Netherlands
2. Department of Neurology, Leiden University Medical Center, Leiden, Netherlands
3. Department of Ophthalmology, Leiden University Medical Center, Leiden, Netherlands
4. Duchenne Center Netherlands



6

The feasibility of quantitative MRI of extra-ocular muscles in myasthenia gravis and Graves' orbitopathy

NMR in Biomedicine, Jan 2021
doi:10.1002/nbm.4407

Kevin R. Keene[1,2] | Luc van Vught[1,3] | Nienke M. van de Velde[2] | Isabeau A. Ciggaar[1,3]
Irene C. Notting[3] | Stijn W. Genders[3] | Jan J.G.M Verschuuren[2,4] | Martijn R. Tannemaat[2]
Hermien E. Kan[1,4] | Jan-Willem M. Beenakker[1,3]

ABSTRACT

Introduction

While quantitative MRI could be instrumental in diagnosis and assessment of disease progression in orbital diseases involving the extra-ocular muscles (EOMs), acquisition can be challenging as EOM are small and prone to eye-motion artefacts. We explored the feasibility of assessing fat fractions (FF), muscle volumes and water T2 ($T2_{water}$) of EOM in healthy controls (HC), myasthenia gravis (MG) and Graves' orbitopathy (GO) patients.

Methods

FF, EOM volumes and $T2_{water}$ values were determined in twelve HC (age 22-65 years), eleven MG (28-71 years) and six GO patients (28-64 years) at 7Tesla using Dixon and multi-echo-spin-echo sequences. The EOM were semi-automatically 3D-segmented by two independent observers. MANOVA and t-tests were used to assess differences in FF, $T2_{water}$ and volume of EOM between groups ($p<0.05$). Bland-Altman limits of agreement (LoA) was used to asses reproducibility of segmentations and Dixon scans.

Results

The scans were well tolerated by all subjects. The bias in FF between the repeated Dixon scans was -0.7% (LoA: ±2.1%), for the different observers, the bias in FF was -0.3% (LoA: ±2.8%) and 0.03cm^3 (LoA: ± 0.36cm^3) for volume. Mean FF of EOM in MG (14.1±1.6%) was higher than in HC (10.4±2.5%). Mean muscle volume was higher in GO (1.2±0.4cm^3) and MG (0.8±0.2cm^3) compared to HC (0.6±0.2cm^3). The average $T2_{water}$ for all EOM was 24.6ms±4.0ms for HC, 24.0ms±4.7ms for MG patients and 27.4ms±4.2ms for the GO patient.

Discussion

Quantitative MRI at 7 Tesla is feasible for measuring FF and muscle volumes of EOM in HC, MG and GO patients. The measured $T2_{water}$ was on average comparable to skeletal muscle, although with higher variation between subjects. The increased FF in the EOM in MG patients suggests that EOM involvement in MG is accompanied by fat replacement. The unexpected EOM volume increase in MG may provide novel insights in underlying pathophysiological processes.

INTRODUCTION

In diseases that involve the extra-ocular muscles (EOM), local muscle weakness or muscle swelling results in complaints such as double vision and ptosis. The auto-immune disease myasthenia gravis (MG) is characterized by fatigable muscles. Auto-antibodies, targeting the acetylcholine receptor (AChR) or muscle specific tyrosine kinase (MuSK), disturb transmission at the neuromuscular junction. In 85% of patients, MG starts with diplopia and ptosis, which are both caused by EOM weakness.[1,227] Currently, MG is diagnosed by either the presence of autoantibodies in serum, or abnormal findings during neurophysiological testing.[5] However, in patients with pure ocular symptoms, the sensitivity of these tests is low (0.44 for testing antibodies, 0.34 for repetitive nerve stimulation and variable between 0.66 and 0.98 for single-fiber EMG).[10] As a result, the diagnosis remains a challenge and there is a need for more sensitive diagnostic tests.[2] Moreover, there is a need for prognostic markers to predict disease progression in ocular MG[228] and refractory MG[8].

Graves' orbitopathy (GO) is an orbital disease caused by auto-antibodies against thyroid proteins, including the thyroid-stimulating hormone receptor (TSH-R).[51,229,230] Due to inflammation and swelling of the EOM, symptoms of GO include upper eyelid retraction, proptosis and diplopia.[65] At a later stage of the disease, fibrosis appears in the EOM as a result of the longstanding inflammation. There is an unmet clinical need to categorize patients, since immunosuppressant medication and radiotherapy are only effective in the active stage of GO.[144]

Quantitative MRI can be used to assess disease progression and to study pathophysiology of skeletal muscle in neuromuscular disease, by assessing fat replacement, muscle size and T2 relaxation time changes.[36] Assessment of such structural changes in individual EOM by quantitative MRI may therefore also contribute to understanding the pathophysiology and pattern of muscle involvement of the EOM in orbital disease.[87,231] As such, it could contribute to identifying diagnostically challenging subgroups in GO and aid in the follow-up of therapeutic response in MG.

In the current work we explored the feasibility and value of quantitative MRI of the EOM in MG and GO patients and healthy controls. These specific patients groups were chosen because distinct structural changes are expected in both diseases and additional diagnostic tests could directly improve clinical care. In MG there is histological evidence for fat replacement, fibrosis and atrophy of the EOM[125], while in GO the EOM are known to be swollen and inflamed at an early stage, and fat-replaced and fibrotic at a later stage of the disease[232]. However, the EOM are challenging to assess with quantitative MRI. Since they are small, the

high resolution needed for MRI of these muscles results in a poor signal-to-noise ratio. In addition, MRI of the EOM is prone to eye-motion artefacts, which reduce the image quality. Moreover, the surrounding bone from the orbit and air from the nasal sinuses lead to an inhomogeneous B0 field. However, recent studies have shown that the increased signal to noise ratio at 7 Tesla, combined with a cued blinking paradigm to reduce eye-motion artefacts and localized shimming to minimize B0 artefacts, enable high resolution imaging of the eyes.[46] [45] Based on these protocols we therefore developed a dedicated protocol for anatomical and functional imaging of the EOM.

METHODS

Participants

All patients were recruited from the outpatient clinics of the neurology and the ophthalmology departments. The diagnosis of MG was confirmed by a neurologist, based on a combination of clinically confirmed fluctuating muscle weakness and the presence of AChR antibodies or abnormal decrement by repetitive nerve stimulation. The diagnosis of GO was confirmed by an ophthalmologist, based on the presence of the characteristic orbital abnormalities (eyelid retraction, redness/swelling of the eyelids/conjunctiva and enlarged EOM or orbital fat) and the presence of anti-TSH-R antibodies. For MG and GO patients, both acute (diagnosis less than 3 months ago) and chronic (diagnosis more than 1 year ago) stages of the disease were included, to assess a wide variety of structural changes to study the sensitivity of the dedicated MRI protocol. Healthy controls were recruited from the Radiology database for healthy controls.

The local medical ethics committee approved the study and all study participants gave written informed consent prior to MR-scanning.

MR examination

We performed chemical shift based water fat separation using the Dixon method to quantify fat replacement and muscle volume and Multi echo Spin-echo (MSE) to assess the T2 relaxation time of water ($T2_{water}$), as an indicator of disease activity[203]. All subjects were scanned in supine position on a 7T Philips Achieva MRI (Philips Healthcare, Best, The Netherlands) using a cued-blinking paradigm[3] with the upper 16 elements of a 32-channel head-coil (Nova Medical). Patients were asked to blink on visual cues provided on a screen; the MRI acquisition was halted during blinking. A 3-point multi-acquisition 3D Dixon scan was acquired (resolution: 0.8x0.8x0.8mm^3, first time to echo (TE)/ΔTE/repetition time (TR)/ flip-angle (FA)/scan time:2.4ms/0.33ms/10ms/3°/3:50min). The 3D volumetric Dixon scan

was planned with the frequency encoding in the left-right direction and was angulated in the coronal plane if needed, to include both orbits in the field of view. In a subset of seven MG patients and five healthy controls, a second Dixon scan was acquired (resolution: 0.7x1.0x0.7 mm^3, first TE/ ΔTE/TR/FA/scan time:2.4ms/0.33ms/10ms/3°/3:50min) to study the value of a higher resolution in the coronal plane, while maintaining the same voxel-volume. These Dixon scans with a higher resolution in the coronal plane were used to study the reproducibility of the fat fraction (FF) measurement, by comparing them to the isotopic Dixon scan. In another subset of participants, containing five MG patients, one GO patient and seven healthy controls, a MSE scan was acquired per orbit (resolution: 1.2x1.2x3.0mm^3, first TE/ΔTE/TR: 12ms/12ms/4000ms, 24 echo's, 3 slices per eye, scan time: 2:44min). This coronal scan was planned with one orbit ins the field of view perpendicular to the medial and lateral rectus muscle. The scan times were kept under four minutes, as previous experience shows this minimizes movement artefacts. Higher order shimming was performed for all scans, to minimize the susceptibility artefacts caused by the bony structures of the orbit and the nasal cavities.

Data analysis and segmentation

Dixon scans were reconstructed with seven-peak reconstruction using the manufacturer's software to yield water and fat images. No T2* correction was performed, given the short echo spacing of 0.33ms. To minimize eye movement artefacts, the scan time was minimized by using a relatively low TR of 10ms. A correction for the resulting T1-weighting in the water and fat images, was incorporated in the FF calculation, assuming T1 values of 1400ms (correction factor: 1.22) and 365ms (correction factor: 1.08) for water and fat respectively[233]:

$$\text{Fat fraction} = \frac{1.08 * \text{Fat}}{1.22 * \text{Water} + 1.08 * \text{Fat}}$$

Scans were excluded if there were obvious motion artifacts. For the Dixon scans, the four rectus EOM (lateral rectus muscle, medial rectus muscle, inferior rectus muscle and the superior rectus and levator palpebrae muscle complex) were semi-automatically 3D-segmented on the water image, using a seed-growing algorithm with a manually set threshold for each subject in ITK-SNAP[234] (Figure 1) by one observer (KK). After automatic segmentation, the regions of interest were manually corrected to remove blood vessels, regions of orbital fat and regions of other eye muscles. The superior rectus muscle was segmented together with the levator palpebrae muscle as one complex, as their proximity hindered accurate separation in some cases. The segmentation was eroded with one acquisition voxel to account for partial volume effects in the FF calculation. To assess the segmentation reproducibility, an independent physician (NvdV) repeated the segmentation process for the muscles of the right eye of all subjects. The average FFs of the EOM were

determined using an in-house developed Matlab script (Matlab 2016a, The Mathworks of Natick, Massachusetts, USA) by averaging the voxels in the regions of interest after exclusion of the eroded voxels and correction of the water-fat-shift between the fat and water image. The EOM volumes were determined by multiplying the number of voxels in the uneroded region of interest with the voxel volume. 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).

The MSE scans were analyzed using a two-component Extended Phase Graphs (EPG) model, consisting of a water and a fat component, as described before.[235] In this model the main scan parameters, such as flip angle and slice profile, were incorporated and the same fixed T1 for the water and fat compartment was used as described above for T1 correction of the Dixon scans. The difference in water-fat shift in the slice direction between the excitation and refocusing pulses, due to different slice selection gradients for both pulses, was also incorporated in the model. This is especially important at high field strengths, as the shift scales linearly with the magnetic field strength and is 1.2mm for the performed MSE scan. A dictionary matching algorithm was used to obtain the T2 of the water component, the B1 map and FF, for each voxel.[41,211] This dictionary was created using $T2_{water}$ values from 10ms to 60ms, $T2_{fat}$ values from 120ms to 200ms, FF values from 0% to 100% and B1 values from 50% to 140%. To obtain a $T2_{fat}$ value for the match, the $T2_{fat}$ is conventionally calibrated on the subcutaneous fat, assuming it consists only of fat tissue and it has a similar T2 as the

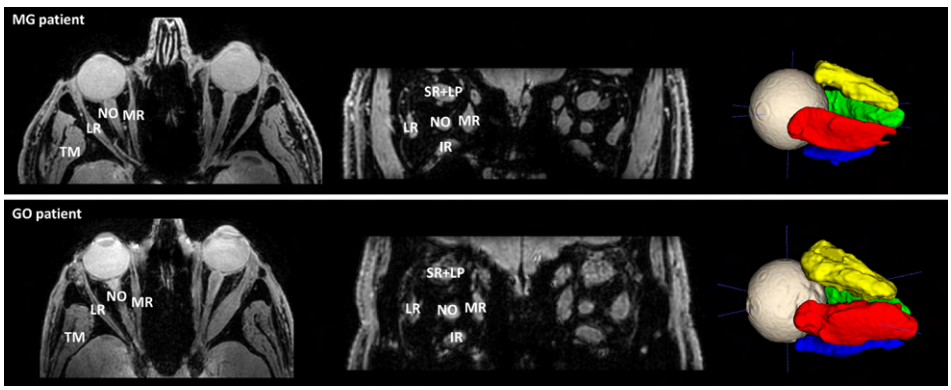


Figure 1. *Transverse slice and a coronal slice of the Dixon water image (left) depicting a myasthenia gravis (MG) patient (top) and Graves' orbitopathy (GO) patient (bottom), showing the lateral rectus muscle (LR), the medial rectus muscle (MR), the inferior rectus muscle (IR) and the superior rectus and levator palpebrae muscle complex (SR+LP). The temporalis muscle (TM) and the optic nerve (NO) are also labeled in the figures. On the right, the 3D regions of interest are shown after automatic segmentation using ITK-SNAP. Notice the enlarged extra-ocular muscles in the GO patient.*

fat replacement in skeletal muscles.[207] However, in this work the $T2_{fat}$ was calibrated on the orbital fat, since subcutaneous fat was not present in the field of view. For the MSE scans, the EOM were segmented by drawing regions of interest per slice using the FF map as an anatomical reference by one observer (KK).

Statistics

The Bland-Altman bias and limits of agreement between the FFs of the two Dixon scans were calculated for all participants together, and separately for the healthy controls and MG patients. The FFs and muscle volumes from the segmentation performed by the two independent observers were compared using Bland-Altman analysis. Mean FF and muscle volumes were compared using MANOVA with post-hoc tests with Bonferroni multiple-testing correction on the patient average and on the individual muscles after averaging the EOM of the left and right eye per patient. Mean $T2_{water}$ values were compared using a two-tailed t-test between healthy controls and MG patients, but not for the GO patient, since only one observation was present. The correlation between age and FF was tested using Pearson correlations for all three groups separately. The correlation between FF and volume was tested for healthy controls for all EOM. Significance level for all tests was set at 0.05.

RESULTS

Subject characteristics

Twelve healthy subjects (41.0±15.9 years, 42% male), eleven MG patients (52.5±15.9 years, 64% male) and six GO patients with active disease (47.7±14.0 years, 50% male, clinical activity score[236] ≥ 3) were included in this pilot study (table 1).

Data quality

All scans but one Dixon scan with higher resolution in the coronal plane from an MG patient were free from visible motion artefacts. This scan was excluded from the analysis. Reconstructed water, FF and $T2_{water}$ maps and the first and last echo of the MSE scan are shown in figure 2.

Agreement on semi-automatic segmentation

The EOM were semi-automatically segmented by two independent observers separately. Figure 3 shows Bland-Altman plots depicting the bias and limits of agreement between the two independent observers for FF (A) and volume (B). The mean bias and limits of agreement in FFs and volume were respectively -0.3% [95% confidence interval(95%CI): -3.1%; 2.6%] and 0.03cm^3 [95%CI: -0.33cm^3; 0.39cm^3].

Table 1. *Baseline characteristics of the three groups*

	Healthy controls N = 12	**Myasthenia gravis N = 11**	**Graves' orbitopathy N = 6**
Age, y	*41.0 ± 15.9*	*52.5 ± 15.9*	*46.7 ± 14.0*
Gender			
Male	*5 (42%)*	*7 (64%)*	*3 (50%)*
Female	*7 (58%)*	*4 (36%)*	*3 (50%)*
Disease onset			
< 3 months	-	*5 (45%)*	*3 (50%)*
> 1 year	-	*6 (55%)*	*3 (50%)*
Medication			
Pyridostigmine	-	*10 (91%)*	-
Prednisone	-	*6 (55%)*	*2 (33%)*
Other immunosuppressants	-	*1 (9%)*	-
MG phenotype			
Ocular	-	*7 (64%)*	-
Generalized	-	*4 (36%)*	-
Anti-TSH-R antibodies	-	-	*6 (100%)*
Anti-AChR antibodies	-	*10 (91%)*	-

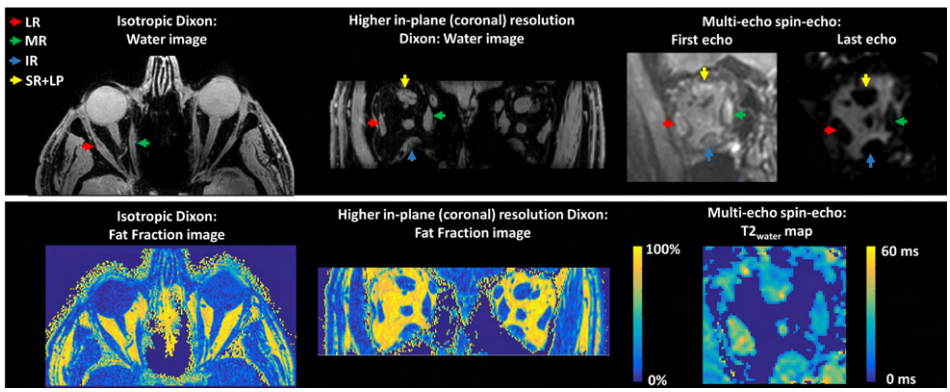


Figure 2. *Examples of water images and fat fraction maps of the isotopic Dixon scan and the Dixon scan with a higher resolution in the coronal plane on the left. On the right the first echo, the last echo and the* $T2_{water}$ *map are shown for the multi-echo spin-echo scan. The colored arrows point out the lateral rectus muscle (LR; red), the medial rectus muscle (MR; green), the inferior rectus muscle (IR; blue) and the superior rectus and levator palpebrae muscle complex (SR+LP; yellow).*

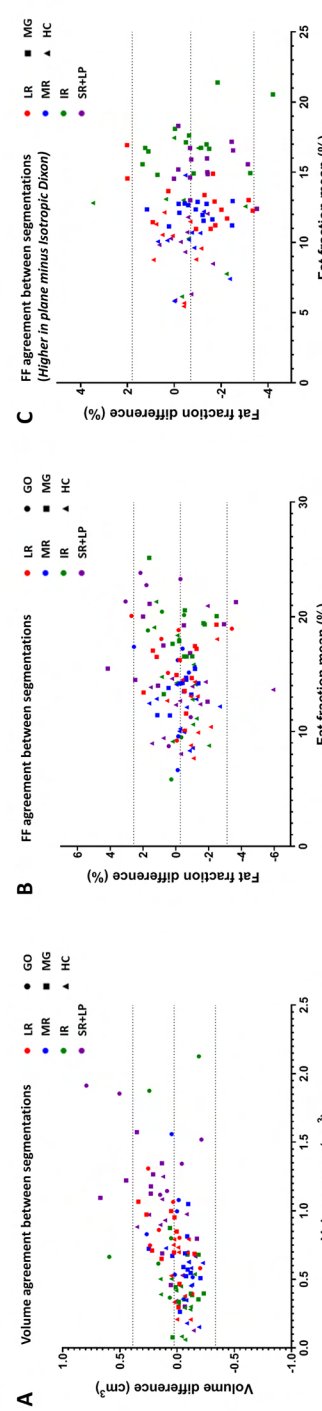


Figure 3. *Bland-Altman plots depicting the agreement between two independent observers who separately semi-automatically segmented the extra-ocular muscles of the right orbits. The four extra-ocular muscles are shown in separate colors: the lateral rectus (LR) muscle, the medial rectus muscle (MR), the inferior rectus muscle (LR) and the superior rectus and levator palpebrae muscle complex (SR+LP). Groups are shown in different shapes: Graves' ophthalmopathy patients (GO), myasthenia gravis patients (MG) and healthy controls (HC). Figure A shows the agreement in muscle volume (bias: 0.03cm^3 [limits of agreement: -0.33cm^3 - 0.39cm^3]). Figure B shows the agreement in fat fractions (bias: -0.3% [limits of agreement: -3.1% - 2.6%]). Figure C shown the agreement between the isotropic Dixon scan and the scan with a higher resolution in the coronal plane (bias: 0.7% [limits of agreement: -3.4%; 1.8%]).*

Reproducibility of fat fraction measurement by Dixon scans

In one patient the second Dixon scan was excluded from the analysis due to severe motion artefacts. On visual inspection, the isotropic Dixon scan was of higher quality compared to the Dixon scan with a higher resolution in the coronal plane. The gain of resolution in the coronal plane was accompanied with a loss of resolution in the transversal plane to maintain sufficient SNR and therefore unwanted partial volume effects were present. The Bland-Altman bias and limits of agreement between the two Dixon scans were -0.7% [95%CI: -3.4%; 1.8%] for all participants, with the Dixon scan with a higher resolution in the coronal plane giving on average lower FFs. In the MG patient group the bias and limits of agreement were -1.0% [95% CI: -3.7%; 1.7%], and in healthy controls the level of agreement was -0.4% [95% CI: -2.6%; 1.7%] (figure 3C).

Muscle volume

The mean muscle volumes of the four recti muscles were higher in GO patients (1.2±0.4cm^3) than in healthy controls (0.6±0.2cm^3; p<0.001) (figure 4). The mean EOM volume was also higher in MG patients (0.8±0.2cm^3) than in healthy controls (p=0.007). Assessing all four muscles separately, the volumes of all four rectus EOM differed significantly between GO patients and healthy controls (lateral rectus muscle: p=0.033, medial rectus muscle: p=0.001, inferior rectus muscle: p<0.001 and superior rectus and levator palpebrae muscle complex: p=0.001). No significant differences were observed comparing individual EOM between healthy controls and MG patients. A difference between GO and MG patients was present in the medial rectus muscle (p=0.035) and the inferior rectus muscle (p=0.001) (figure 5).

Fat fractions

The mean FFs of the EOM where higher in MG patients (14.1±1.6%) compared to those of healthy controls (10.4±2.5%; p=0.008) (figure 6) and similar between MG and GO patients (13.9±4.4%). The variability in FFs appeared higher in GO patients than in MG patients and healthy controls. When analyzed individually, all four rectus EOM differed between MG patients and healthy controls (lateral rectus muscle: p=0.011, medial rectus muscle: p=0.048, inferior rectus muscle: p<0.033 and superior rectus and levator palpebrae muscle complex: p=0.003). A difference between healthy controls and GO patients was observed in the lateral rectus muscle (p=0.029) and the superior rectus and levator palpebrae muscle complex (p=0.003) (figure 7). For each group, Pearson correlations did not indicate a relation between the FF and age (r=0.006 and p=0.86 for healthy controls, r=-0.21 and p=0.69 for GO patients and r=0.31 and p=0.38 for MG patients). In healthy controls no dependence of FF on volume was observed (slope: -2.1 %/cm^3 [95%CI: -4.3 to 0.1], r=0.19 and p=0.06).

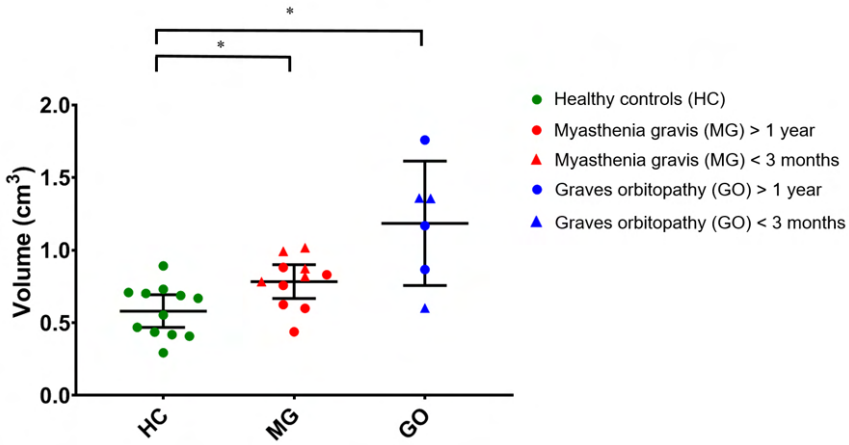


Figure 4. *Mean muscle volumes of the four recti muscles per subject. Data are shown as mean and standard deviation for healthy controls (green), myasthenia gravis (MG) (red) and Graves' orbitopathy (GO) patients (blue). The mean volume was 0.6±0.2cm^3 in healthy controls, 0.8±0.2cm^3 in MG patients and 1.2±0.4cm^3 in GO patients. The MG and GO patients had a significantly higher muscle volume compared to muscle controls (p=0.007 and p<0.001 respectively). Statistical significance is indicated by asterisks.*

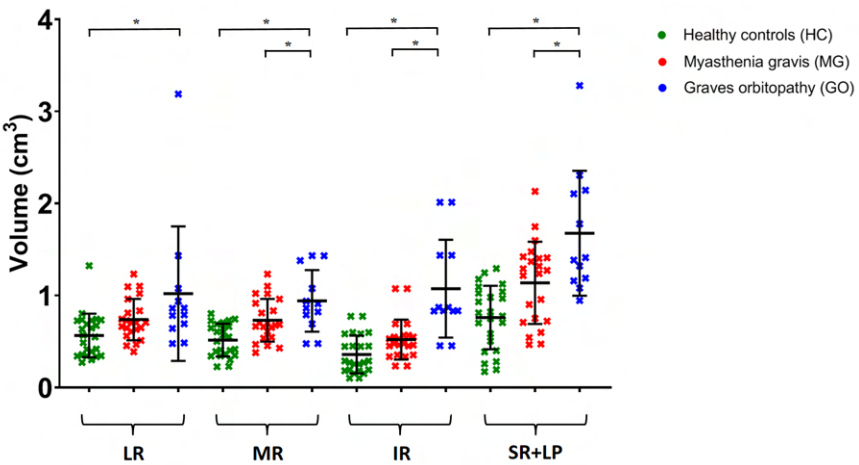


Figure 5. *Muscle volumes for individual extra-ocular muscles of both eyes are shown as mean and standard deviation for healthy controls (green), Myasthenia gravis (MG) (red) and Graves' orbitopathy (GO) patients (blue). All four rectus EOM differed significantly between GO patients and healthy controls (lateral rectus muscle (LR): p=0.033, medial rectus muscle (MR): p=0.001, inferior rectus muscle (IR): p<0.001 and superior rectus and levator palpebrae muscle complex (SR+LP): p=0.001, left and right averaged). The difference between GO and MG patients was present in the medial rectus muscle (p=0.035) and the inferior rectus muscle (p=0.001). Statistical significance is indicated by asterisks.*

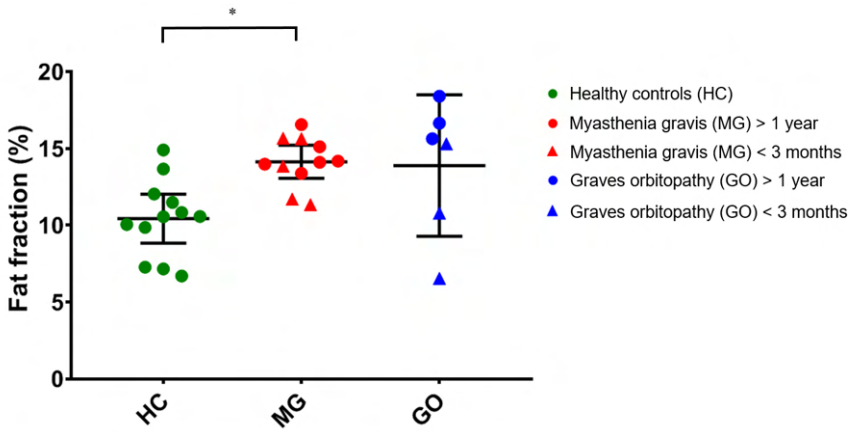


Figure 6. *Mean muscle fat fractions of the four recti muscles per subject, divided per group. Data are shown as mean and standard deviation for healthy controls (green), Myasthenia gravis (MG) (red) and Graves' orbitopathy (GO) patients (blue). The mean fat fraction was 10.4±2.5% in healthy controls, 14.1±1.6% in MG patients and 13.9±4.4% in GO patients. MG patients had a significantly higher fat fraction compared to controls (p = 0.008). Notable is the high variability between muscles in the GO patient group. Statistical significance is indicated by asterisks.*

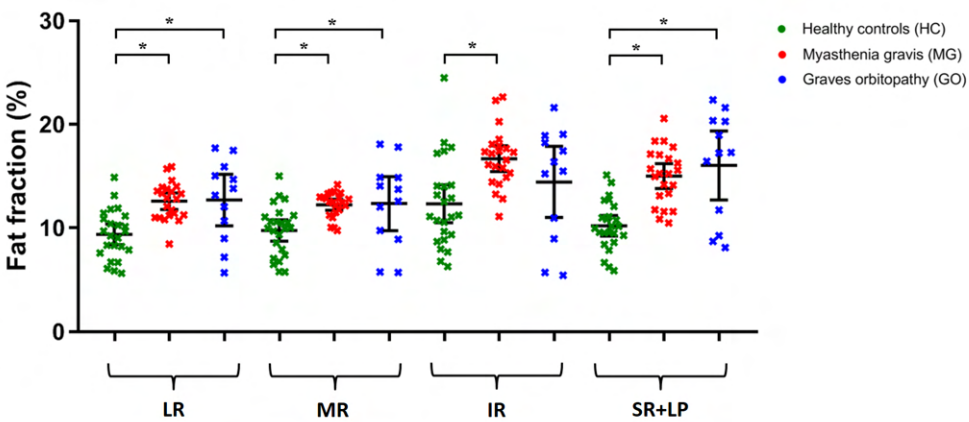


Figure 7. *Fat fractions of individual extra-ocular muscle of both eyes are shown as mean and standard deviation for healthy controls (green), myasthenia gravis (MG) (red) and Graves' orbitopathy (GO) patients (blue). All four rectus EOM differed significantly between MG patients and healthy controls (lateral rectus muscle (LR): p=0.011, medial rectus muscle (MR): p=0.048, inferior rectus muscle (IR): p<0.033 and superior rectus and levator palpebrae muscle complex (SR+LP): p=0.003, left and right averaged). The difference between healthy controls and GO patients was present in the lateral rectus muscle (p=0.029) and the superior rectus muscle (p=0.003). Statistical significance is indicated by asterisks.*

Water T2

The average $T2_{water}$ for all EOM was 24.6±4.0ms for the healthy controls, 24.0±4.7ms for the MG patients and 27.4±4.2ms for the GO patient. No significant differences between healthy controls and MG patients were found. $T2_{water}$ values for individual EOM are depicted in figure 8.

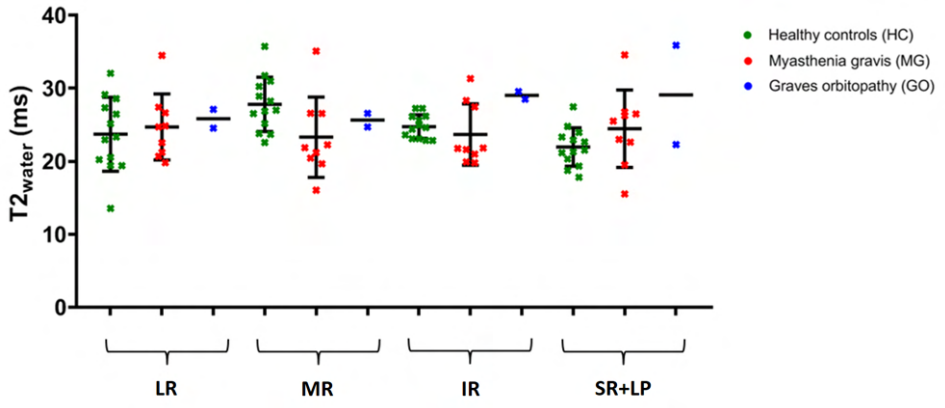


Figure 8. *$T2_{water}$ values of individual extra-ocular muscle of both eyes are shown as mean and standard deviation for healthy controls (green), Myasthenia gravis (MG) (red) and Graves' orbitopathy (GO) patients (blue). For one healthy control and one MG patient, only the EOM from one eye are included. The average $T2_{water}$ for all EOM was 24.6±4.0 ms for the healthy controls, 24.0±4.7 ms for the MG patients and 27.4±4.1 ms for the GO patient.*

DISCUSSION

The present work aimed to study the feasibility of quantitative MRI of EOM in diseases that affect the eye muscles. We showed that the EOM could be reproducibly semi-automatically segmented with a small inter-observer variation and that FFs could be reproducibly measured in healthy controls and MG and GO patients between different scans. With that, we were able to detect small differences in muscle volume and FF between healthy controls and two patient cohorts, as well as $T2_{water}$ values similar to spectroscopy measured values in skeletal muscles, but small differences cannot be excluded as the study was not powered to detect these.
All scans but one were free from visible motion artefacts, indicating that all subjects could adhere to the protocol. The mean bias in FF between the segmentations of two independent observers was low (0.3%) and the limits of agreement were below ±3% in FF, which is in the same range as previously reported values in segmentation of the vastus lateralis in patients with Duchenne muscular dystrophy (bias: 0.1%, limits of agreement: ±1.1%)[38], in healthy

controls and patients with diabetes type 2 in all leg muscle compartments (bias: -0.2, limits of agreement: ±2.6%)[237] and in patients with myotonic dystrophy type 1 in lower leg muscles (limits of agreement: ±2.5%)[238]. The scan-to-scan reproducibility between the two Dixon scans with different resolutions showed a low average bias in FF (-0.7%) and narrow limits of agreement (-3.4% to 1.8%), which is lower than the difference found between the patients and the controls (i.e ~4%), indicating the feasibility of assessing FFs using this method to detect affected muscles. For comparison, Dixon measurements in skeletal muscles have been reported to have a reproducibility (limits of agreements) of ±1% for water-fat phantoms[239], ±1.1% for thigh muscles in healthy controls[240] and ±2.5% for the leg muscles in patients with Duchenne muscular dystrophy.[200] These techniques could be used in future longitudinal studies to study the disease progression in conditions affecting the EOM and to relate time varying clinical observations such as painful eyes to the current condition of the EOM in terms of fat replacement and inflammation.

We found FFs of approximately 10% in the EOM, which is higher than the 2-5% commonly observed in skeletal muscles in healthy volunteers.[240–242] Although the high FF could be an intrinsic property of EOM, this could also be the result of noise bias. Earlier studies in skeletal muscles have shown that low SNR could result in an overestimation of the FF in the low FF ranges.[243] Although the strong contrast between the orbital fat and the muscles and globe, suggests more than sufficient SNR for anatomical evaluation of the images, some noise can be observed in the temporal muscle, a large skeletal muscle close to the orbit. In a representative healthy subject, this temporal muscle also showed a relatively high FF of 9.3% (data not shown). A Dixon-reconstruction of fourfold down-sampled source images of the same subject, which increased the SNR twofold, resulted in a reduction in FF from 9.3% to 5.1% was observed, showing that the used protocol is susceptible to noise bias in the low FF ranges. This could explain the relatively high FF observed in the EOM. We therefore recommend that future studies use a higher SNR for the Dixon acquisition, for example by using a higher flip angle or a two flip angle approach.[243] Additionally, the overestimation of the FF could have been caused by partial volume effects of the EOM with the surrounding orbital fat, due to the small size of these muscles (~ 1cm^3).

Muscle volumes were measured using 3D segmentation on the water image from the Dixon scans. The limits of agreement for the muscle volumes segmented by the two independent observers ranged between -0.33cm^3 and 0.39cm^3. These limits do not reach the 0.6cm^3 difference observed between GO patients and healthy controls, but are higher than the difference of 0.2cm^3 observed between the MG patients and healthy controls. On visual inspection (figure 3), the highest differences between observers were found in the superior rectus and levator palpebrae muscle complex, which is the hardest to delineate, because the

complex contains two separate muscles. An overall dependence of volume difference between observers on mean volume was observed (figure 3A), but there was virtually no mean bias (0.03cm^3). Moreover, as no dependence of FF on the volume was present in healthy controls, the segmentation was not likely to have influenced the observed FF differences between groups.

Using MSE and EPG analysis, we found T2 values of the myocytic compartment ($T2_{water}$) for EOM to be comparable to the $T2_{water}$ in healthy skeletal muscle, as measured with MR spectroscopy.[39,222,244] The standard deviation of $T2_{water}$ between the EOM (4.0ms, 4.7ms and 4.2ms for respectively healthy controls, MG patients and GO patients) was higher than previously measured using an EPG model in leg muscles between subjects (1.4ms-2.1ms)[41,241]. This variation between $T2_{water}$ of muscles could be expected in the patients as different patients have a different pattern of muscle involvement. In GO, for example the inferior rectus muscle is usually affected earlier in the disease, followed later by the other EOM.[87] In the healthy controls, however, a smaller variation, comparable to other skeletal muscles, was expected. Two assumptions in the analysis may have contributed to this higher variation. Firstly, the calibration of the $T2_{fat}$ was performed on the orbital fat, under the assumption that its composition is similar to intra-muscular fat[216]. Orbital fat is, however, known to be involved in inflammatory conditions, such as GO.[245] This variation in orbital fat T2, which is not necessarily also present in the T2 of the intra-muscular fat, might result in an erroneous variation in intra-muscular $T2_{fat}$. Secondly, our model assumes that the ROI to calibrate the $T2_{fat}$ only contains orbital fat. However, in reality there is vascularization in the orbital fat which will result in a small, patient-specific, inaccuracy in the measured orbital $T2_{fat}$, which will subsequently influence the obtained $T2_{water}$. A high variation between individual EOM in patients is to be expected, as MG and GO are known to selectively affect specific EOM.[78,90] Due to the high variation, the power was limited to detect mean differences in $T2_{water}$ between groups. For future studies, we recommend establishing baseline values per EOM and use Z-scores to detect EOM with an abnormal $T2_{water}$.

Although the patient cohorts were small, clear differences were found between cohorts. The measured FFs of the EOM were higher in MG and GO patients compared to healthy controls and muscle volumes were higher in GO patients. Edematous EOM are a common finding in GO patients[65] and fat replacement of EOM has been described in late-stage GO patients[51]. The variability in muscle volume was high in GO patients, possibly due to the inclusion of patients with varying disease severity and the inflammation of only a portion of the EOM in individual patients.

The increase in FF in the EOM in MG has been described in histological studies[125], but the increase in muscle volume in the EOM of MG patients was unexpected. In MG, AChR or MuSK antibodies can result in structural damage of the muscle membrane at the neuromuscular synapse, which is usually located at the middle of the muscle fiber, but this is not a dystrophic process.[2] In dystrophic muscle diseases fat replacement of muscle tissue is commonly observed, and can almost completely replace normal muscle tissue.[37,246] In neuropathic diseases, like hereditary motor and sensory neuropathy (HMSN) or hereditary neuropathy with liability to pressure palsies, FFs have been observed up to 15% in calf muscle[124]. Which is a relatively small increase (up to 10%) in FF, comparable to the small increase in FF (up to 8%) that we observed in the EOM in MG. In a recent case report, histological findings in a seronegative MG patient also included adipocyte replacement of muscle fiber,[247] and fat replacement has been seen in bulbar muscles of MuSK positive MG patients.[18] Our results agree with these observations, and indicate that involvement of EOM in MG patients is accompanied by fat replacement. Interestingly, our results so far suggest that FFs of MG patients with a disease onset less than three months before the scan did not appear to differ from chronic MG patients.

Severe MG can induce chronic denervation and therefore might behave as a neuropathic condition. Muscular atrophy in skeletal muscle is observed in 10% of AChR+ MG patients with long-standing disease.[20] In MuSK+ MG patients early atrophy is seen in the bulbar muscles.[18,19,248] In other neurogenic muscle diseases, like HMSN, also a decrease in muscle volume has been observed.[124] However, in this work we observed an increase in EOM volume in MG patients, which is not consistent with the expected atrophy. This unexpected increase in muscle volume raises questions about its pathophysiology. Firstly, it is possible that inflammatory processes in the EOM result in swollen muscles. In MG, complement activation is observed, resulting in activation of the membrane attack complex and destruction of the neuromuscular endplate. [249] To what extent this inflammatory process affects the EOM is unknown. Secondly, (pseudo)hypertrophy may be an explanation of the increase in EOM volume, which had been observed in some muscular dystrophies.[250]

A limitation of this work was that the groups in this pilot study were not fully age matched, as the maximum age of healthy volunteers was limited to 65 years, based on the inclusion criteria of our IRB approval. This upper age limit did not apply to both patient groups. However, an earlier study showed that the EOM do not change in muscle thickness, size and fatty tissue volume between age groups.[251] In addition, when analysing all three groups in this study separately, we did not see any age effects on EOM FFs. Therefore, age does not appear to be a major confounder in our study. Secondly, due to the exploratory nature of this pilot study with ongoing sequence development, not all sequences were applied to all patients.

As a result, the second Dixon scan for reproducibility and the MSE sequence was only performed in a subgroup of patients, and the T2 relaxometry only to one GO patient. Lastly, this study was performed on a 7 Tesla MR-scanner to increase the SNR at the high resolution required for imaging the EOM.[46] However, downsides of scanning at 7 Tesla are the increase in susceptibility differences, B0 inhomogeneities and chemical shift artefacts. These challenges were successfully solved using higher order shimming and the in-plane chemical shift displacement was set at 1.0 acquisition pixels and corrected for in postprocessing. Future studies can determine if these same results can be achieved at clinical field strengths, such as 3 Tesla.

In conclusion, despite the challenges of size and movement, quantitative MRI is feasible for measuring FFs and muscle volumes of individual EOM in healthy controls, MG and GO patients at 7 Tesla. We measured water T2 relaxation times comparable to values measured in skeletal muscle, but found a high variability between individuals and between different EOM. Detection of group differences for this parameter will therefore only be possible when large differences between groups are present, only affected EOM are included in the analyses, or when using large sample sizes. We found an increase in EOM volume and FF in MG and GO patients. The unexpected increase in EOM volume in MG may provide novel insights in underlying pathophysiological processes. The diagnostic value of these findings and their value in the follow-up of disease progression should be the subject of future prospective studies.