



T2-relaxation mapping and fat fraction assessment to objectively quantify clinical activity in thyroid eye disease: an initial feasibility study

Tilak Das ¹ · Jonathan C. P. Roos^{2,3} · Andrew J. Patterson¹ · Martin J. Graves¹ · Rachna Murthy^{2,3}

Received: 7 November 2018 / Accepted: 16 November 2018 / Published online: 11 December 2018
© The Royal College of Ophthalmologists 2018

Abstract

Imaging in thyroid eye disease (TED) is used to exclude other diagnoses, assess for apical crowding and plan surgery. But to quantify TED activity objectively, subjective clinical scoring assessments remain the norm. Magnetic resonance imaging (MRI) T2-relaxation times correlate with extra-ocular muscle (EOM) inflammation, but are confounded by signal from fat. We investigated whether T2-relaxation mapping in combination with fat fraction (FF) measurements could quantify disease activity in EOMs objectively. Sixty-two TED patients and six controls were enrolled for coronal short tau inversion recovery (STIR), T2 multi-echo fast-spin echo and multi-echo fast-gradient echo MRI of the orbits. STIR signal intensity ratios (SIRs), T2-relaxation times and percentage FF were derived for inferior, lateral, superior and medial recti bilaterally. Twelve patients were re-scanned following immunosuppressive treatment. The results found a positive correlation for all subjects between T2 and SIR ($p < 0.001$), but only mean T2 differed significantly between patients and controls ($p < 0.001$). We measured FF in EOMs for the first time and found it greater in TED ($p < 0.001$). There was also a significant reduction in mean T2 after treatment, with a corresponding reduction in the clinical activity score (CAS) in almost all patients. We show that T2-relaxation times differentiate between normal and inflamed EOMs and are responsive to treatment. Combined, uniquely, with FF measurement in EOMs, an objective, quantitative marker of inflammation in TED-affected muscles could be derived. T2-relaxation times mirrored improvements in CAS after treatment, occasionally preceding them. Rarely, they diverged, suggesting limitations in the CAS as a disease burden marker.

Introduction

Assessing TED and the role of imaging

A wide array of imaging techniques, summarised in Table 1, have been proposed to aid in the diagnosis and management of thyroid eye disease (TED) [1]. While most TED patients may not require any imaging, it can be necessary for assessment of orbital apex crowding affecting

the optic nerve, planning decompression surgery and to exclude other orbital diseases. This is particularly so with asymmetrical presentations when myositis, lymphoma, carotico-cavernous fistulae, varices, IgG4 disease, meningiomas and other malignancies should be considered [2].

A further and emerging role for orbital imaging is to help guide and determine response to treatment. Several principally magnetic resonance imaging (MRI)-based methods have been developed in an effort to quantify disease activity objectively (see Table 2). However, they have drawbacks, including confounding by surrounding tissue effects, the use of semi-quantitative relative values with a lack of day-to-day reproducibility and complexity of sequences required to derive absolute values.

Given these limitations, clinicians still rely principally on subjective scoring systems to quantify disease severity and activity clinically; a widely used tool is the clinical activity score (CAS) [3], but it has been criticised as too binary, subjective and lacking in sensitivity [4]. It is also poorly designed for monitoring of improvement; [5] even marked

✉ Tilak Das
tilak.das@addenbrookes.nhs.uk

¹ Department of Radiology, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK

² Department of Ophthalmology, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK

³ Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

Table 1 Imaging modalities for thyroid orbitopathy and their characteristics

Imaging modality	Type of signal	Dangers	Utility in TED	Advantages	Disadvantages	Provides a quantifiable marker of disease activity?
CT	X-ray	Radiation- and iodine-based contrast	Better than MRI at identifying enlarged muscles [23]. Density of muscles can correlate with disease [24]	Fast. Better bone resolution than MRI. Good for assessment of apical crowding [1]	Radiation. Volumetry of muscles may not correlate with disease activity [25]	Mainly proptosis
MRI	Nuclear magnetic resonance	Patients with certain active or passive implants may be contraindicated	Better than CT at identifying areas of inflammation in muscle	Better contrast between soft tissues than CT. Different sequences provide specific anatomical or physiological information	Slow. Acoustic noise. Costly. Movement artefacts or semi-quantitative data	Potentially, but most clinical protocols provide qualitative or semi-quantitative data
Ultrasound	Sound echo	None	Limited	Rapid. Available. Can perhaps exclude scleritis and intraocular pathology	User-dependent. Limited depth. No visualisation of apex. Limited reproducibility	No
Doppler	Sound echo	None	Limited clinical utility but blood flow in superior ophthalmic vein is reduced in TED [26]	Rapid. Available	Difficult to perform. Inter-observer variability. Not specific for TED	No
Octrooscan with ¹¹¹ In	γ -Ray scintigraphy	High radiation	Limited	Orbital uptake of this labelled somatostatin analogue is greater in TED	Non-specific. High cost. Poor availability. Requires careful standardisation	No
Octrooscan with ^{99m} Tc	γ -Ray scintigraphy	Lower radiation	Limited. But significant correlation was found between CAS and the orbital uptake. Identifies active disease	Lower cost. Greater availability. Higher energy so improved resolution. Shorter acquisition time	Radiation. Invasive	No
Gallium-67 scintigraphy	γ -Ray scintigraphy	Radiation	Limited. Able to detect response to treatment	Equivalent positive predictive value to octreotide and T2-relaxation time on MRI	Invasive. Requires careful standardisation	No, too difficult
FDG-PET/CT	Positron scintigraphy	Radiation	Limited but able to identify activity even when MRI normal [27]	May detect early/subclinical disease	Radiation. Not widely used	No
Thermal	Heat	None	Limited. But higher temperatures recorded in TED	Non-invasive. Changes in response to IV steroids [28]	Low specificity. Requires complex equipment to measure accurately	No, too much variability

Table 2 MRI-based techniques used for TED assessment

MRI sequence	What it is	Utility in TED	Advantages or disadvantages	Provides a quantifiable marker for TED?	Reference
T1	A predominantly anatomical scan	Muscle size measurement. Detection of fat (high signal) within muscle (lower signal)	No measure of disease activity. Non-quantitative other than anatomical measurements	Muscle dimensions (usually 2-dimensional). T1 signal interpretation is qualitative	
T1 + Gd	A scan sensitive to abnormal contrast enhancement with gadolinium	Inflamed muscles demonstrate greater enhancement	Requires IV contrast medium. Similar information can be obtained from non-contrast sequences such as STIR	No (qualitative due to variability)	
T1 + Gd SIR	A ratio calculated by comparing signal in a muscle after contrast with a non-enhancing structure (e.g. thalamus) [16]	Differentiates active disease from inactive and healthy. Allows tracking change over time	Signal correlates with CAS [29]. Requires IV contrast medium. Does not differentiate inactive disease from healthy controls	Semi-quantitative, using a ratio with an arbitrary denominator	Politi et al. [19] Tortora et al. [29]
T2	A scan to mainly identify water content (generally higher intensity a.k.a. longer T2-relaxation time suggests higher water or fat content)	Can help detect activity due to higher signal with oedema and inflammation	Standard protocol for orbital imaging but not for TED. Normal muscle is dark on a bright background of orbital fat. Inflamed muscle may be bright and poorly differentiated from surrounding fat	No (qualitative due to variability)	
T2 mapping	Scan to derive quantitative measurements of T2-relaxation time	Can detect activity and differentiate healthy from active disease. Correlates with response to treatment	T2-relaxation time correlates with CAS. Allows tracking change over time. Should be independent of equipment. Sequence may not be readily available	Yes, but possible confounding by signal from fat	Ohnishi et al. [16], Mancini et al. [15]
STIR sequence	A T2-weighted sequence with suppression of fat signal	Highlights inflammation in muscle and helps to delineate active phase	Standard readily available sequence. Signal correlates with CAS [12, 30]. Susceptible to artefact (poor fat suppression, motion, paranasal sinus air, braces, etc.)	No (qualitative due to variability)	Hoh et al. [31] Mayer et al. [14] Tortora et al. [12]
STIR SIR	Signal intensity ratio comparing an unaffected muscle with an affected muscle or other structure. For example, EOM and temporalis or normal white matter	Can help detect activity and differentiate healthy from active disease. Allows tracking change over time	Signal correlates with CAS but does not reliably differentiate TED patients from healthy controls. Variability with equipment	Semi-quantitative, using a ratio with an arbitrary denominator	Politi et al. [19]
T2 with fat suppression (e.g. Dixon method)	Aside from STIR, numerous other techniques are available to suppress fat signal. The Dixon method is a specific MRI technique that allows separate 'water' and 'fat' images to be generated	Water fraction correlates with activity scores and can monitor effects of therapy	Exact methods vary between manufacturers. Most are untested for EOMs	Quantitative evaluation of water content and fat fraction in orbital fat. Untested for extra-ocular muscles	Kaichi et al. [32]
Fat fraction mapping	Technique that allows quantification of fat fraction by exploiting differences in physical properties of protons in water vs. lipid environments	Quantitative evaluation of fat content within extra-ocular muscles	Allows tracking over time. Sequence may not be readily available	Yes. Fat fraction shows demonstrable differences in EOMs between patients with TED and healthy controls	This report

Table 2 (continued)

MRI sequence	What it is	Utility in TED	Advantages or disadvantages	Provides a quantifiable marker for TED?	Reference
DWI	Scan that is sensitive to diffusivity of protons. Generally, oedema results in greater extra-cellular space and increased diffusivity. ADC is a marker of this	Differentiates between patients with TED (higher diffusivity in EOMs) and healthy controls	Helpful for excluding lymphoma and myositis [1]. Signal correlates with CAS but values do not differ between patients with active and inactive disease [19]. Susceptible to artefact and distortions at air–bone interfaces. Requires optimisation for use in orbits	ADC is a quantitative measure. Reproducibility is not established	Politi et al. [19] Lingam et al. [30]
DCE-MRI	Serial scanning of a region of interest after injecting IV contrast medium to look at dynamics of contrast enhancement	Highlights changes in microcirculation in TED. Can correlate with change from active to fibrotic stage	Signal correlates with TED activity. Requires contrast and careful timing for image acquisition. Post-processing and analysis is complex	Quantitative. However, relationship between contrast dynamics and TED EOM composition is unclear [33]	Sun et al. [34] Jiang et al. [35] Guo et al. [36] Taoka et al. [33]

STIR short tau inversion recovery, *DWI* diffusion-weighted imaging, *DCE-MRI* dynamic contrast-enhanced MRI, *Gd* gadolinium, *ADC* apparent diffusion coefficient

amelioration of any component of the CAS will not improve the actual score unless the feature resolves completely [6]. Not surprisingly, more complex scoring systems have been introduced such as the VISA [7], modified NOSPECS [8] and EUGOGO (see www.eugogo.eu) scoring tools. But fundamentally they remain subjective and are cumbersome to administer [7].

Other proposed methods for assessing disease progression include serial binocular single vision mapping, measurements of proptosis [9] or Hess charts [10]. But these may not reflect activity or severity dynamically as the active inflammatory phase of TED can be followed by a chronic scarring phase with a static effect on ocular motility. Serial thyroid-stimulating hormone receptor antibody levels have been postulated to provide a further objective and quantitative means by which to assess thyroid eye disease activity but have not yet been widely adopted [11].

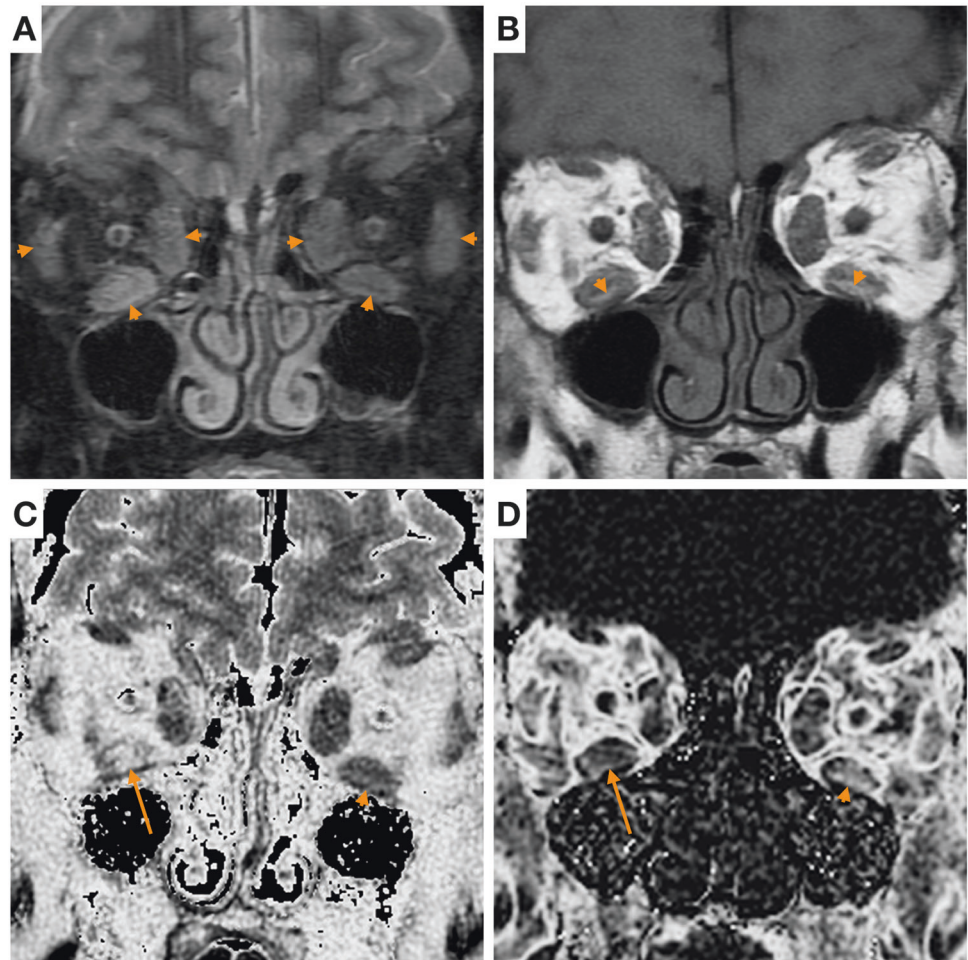
Proposed quantification of TED by modified MRI protocol

There has been great interest in using MRI to quantify disease activity, severity and response to treatment in TED. Some of these are summarised in Table 2. However, there is inherent variability in MRI measurement, influenced by technical factors such as variability in scanners, sequences used and physiological factors such as changes over time (from diurnal to age related) or change in disease activity. Signal intensities seen on images are not usually directly proportional to specific tissue properties such as density or molecular composition. This variability in measured MR signal values makes them unsuitable for quantifiable comparison between scans—even within an individual—unless sequences are specifically designed for quantification. To get around this, groups have used signal intensity ratios (SIRs) from standard clinical images such as those generated from STIR (short tau inversion recovery) sequences. These are T2-weighted fat-suppressed images in which higher intensities reflect higher water content. EOM intensity is compared with other unaffected muscles or even non-muscular tissues [12–14], but such ratios are less sensitive. Mathematically, small variabilities can be amplified to significantly alter these ratios.

Though the MR signal can be variable, the decay rate (or relaxation time) of the signal is not, being a physical property of the tissue that can be expressed numerically. T2 (or transverse)-relaxation times depend strongly on water content of tissue and longer T2-relaxation times correlate with greater EOM inflammation [15, 16]. However, signal from fat can confound these quantitative measurements as these also represent a variable combination of water and fat.

Fat fraction (FF) mapping is a method that takes advantage of the slightly different resonant frequencies of

Fig. 1 Images of the orbits of a patient with thyroid eye disease. **a** Coronal short tau inversion recovery (STIR): T2-weighted fat-suppressed image showing enlargement and heterogeneous hyperintensity of extra-ocular muscles (arrowheads). **b** Coronal T1-weighted image: areas of T1 hyperintensity within the muscles are consistent with fatty change (arrowheads). **c** T2-relaxation map shows areas of longer T2-relaxation (higher intensity) in the right inferior rectus muscle (arrow) compared to the left (arrowhead). **d** Fat fraction map shows higher fat fraction (higher intensity) in the left inferior rectus muscle (arrowhead) compared to the right (arrow)



water and fat to generate images that separate the contribution of water and fat to the signal, allowing quantitative measurements of the FF within a tissue [17, 18].

In this study, we sought to compare T2-relaxation mapping to STIR SIR and, for the first time, to calculate FFs for EOMs. The combination of FF data with T2-relaxation times may refine signal quantification and could then represent a robust quantitative measure of disease activity.

Methods

Between 2014 and 2016, 62 patients with TED and six healthy controls were recruited for this study which had received ethical approval from our hospital internal ethics review board.

All patients received a full ophthalmological work-up, including determination of CAS and blood tests. MRI of the orbits was performed on a 1.5 T MRI scanner (450, GE Healthcare, Waukesha, WI, USA) including coronal STIR, T2-weighted multi-echo fast-spin echo and multi-echo

fast-gradient echo sequences [18] (Fig. 1). The sequences for T2 relaxation and FF mapping took an additional 10 min in the scanner to acquire.

Twelve patients who required active management with immunosuppression with cyclosporine and prednisolone were re-scanned following their treatment.

STIR SIRs, T2-relaxation times and percentage FF were measured by a neuroradiologist with more than 7 years' experience, blinded to the clinical data, from regions of interest in the muscle belly of inferior, lateral, superior and medial rectus muscles bilaterally in all individuals. Thereafter, Pearson's correlation between T2 and SIR was calculated for each individual muscle. Unpaired *T* tests were performed to investigate if T2, SIR and FF were different between the thyroid eye disease patients and normal. Paired *T* tests were performed to investigate if T2 and SIR were different in patients receiving immunosuppressive treatment. For FF comparisons, to meet the normality assumptions for standard parametric tests, the distributions were log transformed. Statistical analysis was performed in R (The R foundation for statistical computing, Vienna, Austria).

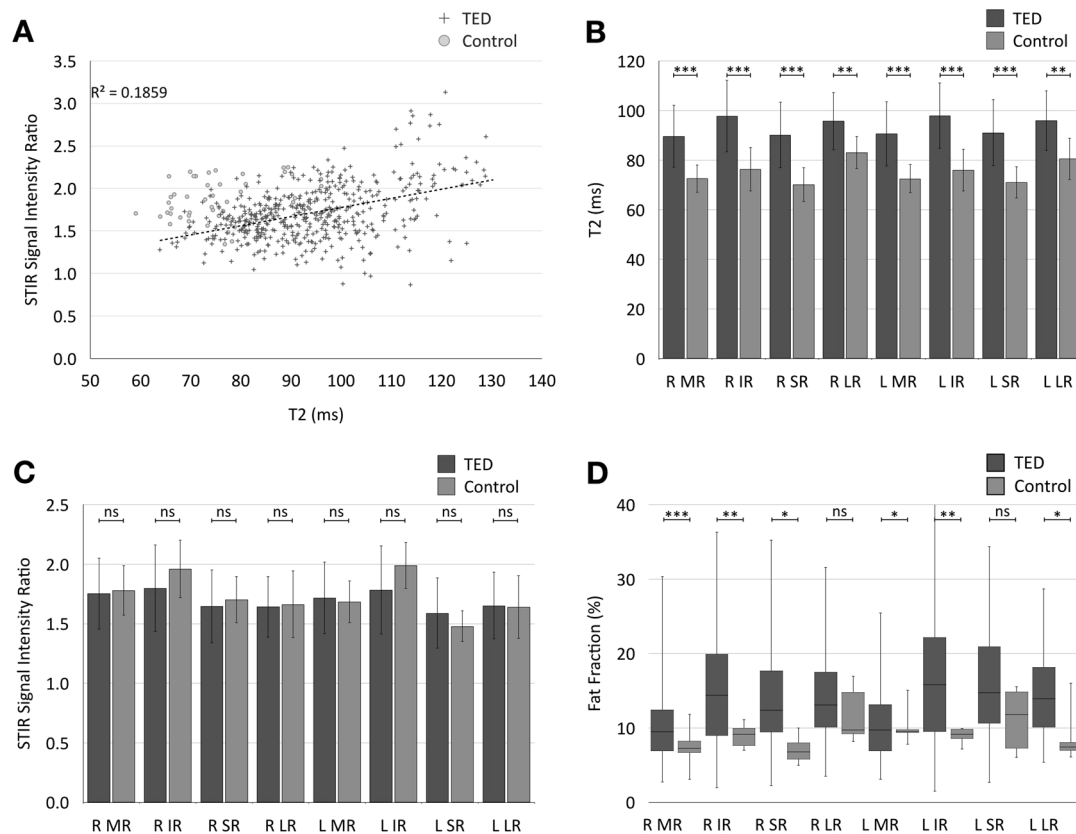


Fig. 2 **a** Scatter plot showing a positive correlation between T2-relaxation times and STIR SIR in extra-ocular muscles in TED patients and healthy controls ($r = 0.42$ in TED patients). However, the groups are poorly separated across the SIR axis and better separated by T2-relaxation time. **b** Comparison of mean (\pm SD) T2-relaxation times for individual extra-ocular muscles across the TED and control groups. T2-relaxation time is higher in the TED patients for each extra-ocular muscle. **c** Comparison of mean (\pm SD) SIR for individual extra-ocular muscles across the TED and control groups. There was no statistically

significant difference in mean SIR of any of the measured extra-ocular muscles between the two groups. **d** Boxplot showing median and range of fat fractions for each extra-ocular muscle compared between TED and control groups. There were significant differences detected in some of the extra-ocular muscles. Mean FF for all muscles as a group was significantly different between TED and control groups ($p < 0.001$). (R = right, L = left, MR = medial Rectus, IR = inferior rectus, SR = superior rectus, LR = lateral rectus. For p values: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns non-significant.)

Results

We initially investigated if the T2-relaxation data correlated with SIR. As shown in Fig. 2a, for both TED and normal groups, there was a positive correlation between T2 and SIR ($r = 0.43$, $p < 0.001$).

Mean T2 for each rectus muscle category differed significantly between the TED and normal groups (combined mean T2 TED: 93.92 ± 12.80 ms (95% CI: 92.74–95.11); normal: 75.30 ± 7.96 ms (95% CI: 73.20–78.25), $p < 0.001$; Fig. 2b). However, there was no significant difference in mean SIR (combined mean SIR TED: 1.71 ± 0.32 (95% CI: 1.68–1.74); normal: 1.74 ± 0.25 (95% CI: 1.66–1.82), $p = 0.52$; Fig. 2c).

Overall mean FF was significantly greater in TED patients than in healthy volunteers (mean FF TED: $13.98 \pm 7.0\%$ (95% CI: 13.33–14.64), normal: $9.27 \pm 3.14\%$ (95% CI: 8.38–10.15), $p < 0.001$). There was substantial variability in the mean FF for each muscle in the patient group (Fig. 2d).

As shown in Fig. 3, there was a significant difference in mean T2 as well as SIR between pre- and post-treatment examinations in 12 TED patients receiving treatment with immunosuppression (mean difference T2: 12.05 ± 17.79 ms, $p < 0.001$, mean difference SIR: 0.26 ± 0.56 , $p = 0.001$). On an individual basis (Fig. 3c), a reduction in muscle T2 was related to a reduction in CAS in patients.

Discussion

In this feasibility study, we sought to evaluate if existing MRI protocols can measure muscle inflammation and disease activity in TED patients. We show that measurement of T2-relaxation time is a more robust method for assessing disease activity than STIR SIR. As fat is implicated in the disease process in TED and is a potential confounder to the measurement of T2-relaxation times, we also describe the novel measurement of FF in EOMs.

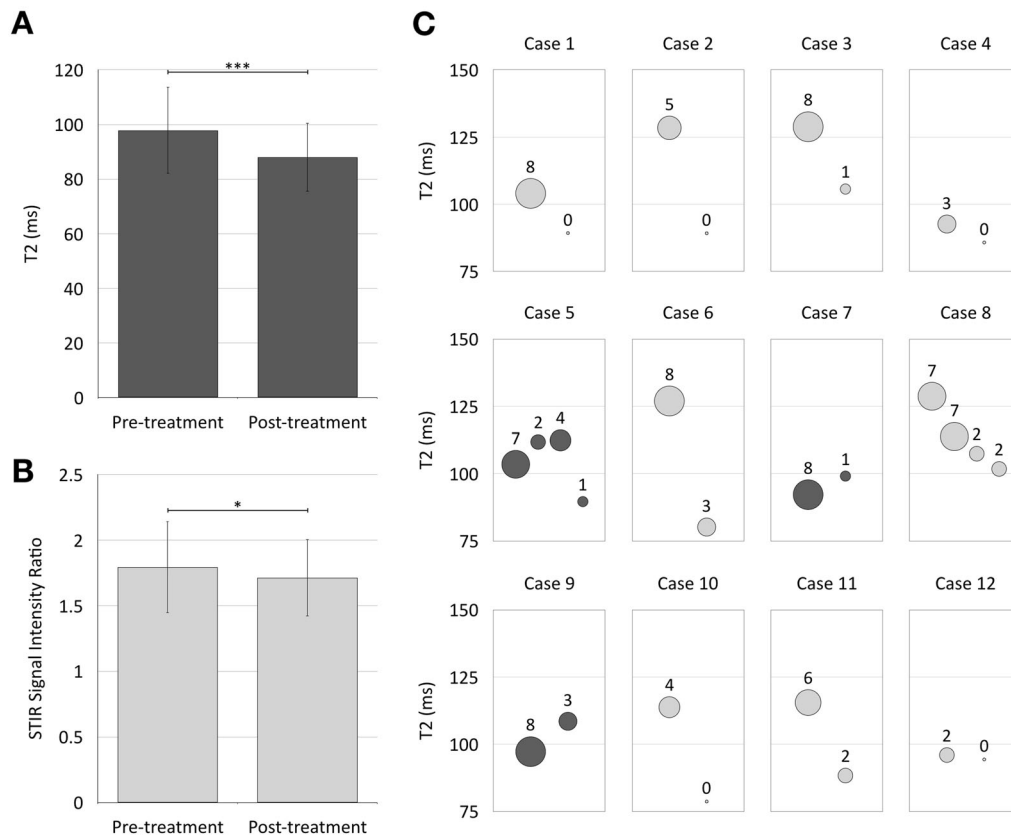


Fig. 3 The mean T2-relaxation time (**a**) and SIR (**b**) in extra-ocular muscles of pre-treatment baseline examinations was higher than on post-treatment examinations (error bars—standard deviation; **a** *** $p < 0.001$; **b** * $p = 0.04$). **c** shows bubble plots of 12 individual cases with pre- and post-treatment examinations. Serial examinations are represented on the x -axis in arbitrary units of time. Maximal T2-relaxation times are plotted in the y -axis with CAS represented by the

area of the bubbles (annotated for each data point). In most cases, T2-relaxation time for the worst affected muscle correlated with CAS, with both reducing following treatment. However, in cases 5, 7 and 9 (highlighted by darker shade), there is initially an increase in T2-relaxation time despite a decrease in recorded CAS. In case 8, there is initially a decrease in T2-relaxation time before a decrease in CAS is detected (see text for further discussion)

T2-relaxation times could clearly differentiate abnormal from normal populations of EOMs, as well as pre-treatment from post-treatment status, in turn correlating with CAS. The correlation between T2-relaxation time and CAS as well as post-treatment status has been investigated previously [15, 16]. However, the significant difference demonstrated in this study between TED patients and controls allows normative values to be considered. T2-relaxation time appears to be a reliable measure of disease activity, allowing the assessment of individual muscles.

Although T2-relaxation time does correlate with SIR, which has also previously been shown to correlate with CAS and post-treatment status [12, 19, 20], the lack of a statistically significant difference in SIR between TED patients and controls at the level of individual muscles is striking and suggests that SIR is less discriminatory and unsuitable for quantitative assessment of disease activity. SIR measurements suffer from the need to choose an appropriate but arbitrary denominator, such as signal intensity from temporalis muscle or normal appearing white

matter on the same image as the measured EOM. T2-relaxation time, as an estimate of an absolute value, is independent of machine and sequence-specific parameters that may affect SIR measurements and a more robust measure for application in longitudinal and cross-sectional comparisons.

In the temporal correlation between maximal T2-relaxation times and CAS (Fig. 3c), there were at least three cases with contradictory assessments (highlighted as darker shade). In each of these, there is a decrease in CAS despite an apparent increase in T2-relaxation time that suggested increased disease activity. Interestingly, clinical follow-up supported the radiological assessments, suggesting that T2 relaxation is a more sensitive marker of inflammation and more dynamic assessment of change. Specifically, in case 5, MRI picked up disease activity before it was reflected in the CAS. In case 7, the MRI results are seen to contradict the high initial CAS—however, this was a falsely positive high CAS due to anterior segment inflammation, not due to TED. Finally, in case 9,

the MRI detected myositis before it became apparent on the CAS. In addition, case 8 in the series demonstrates an overall correlation between T2 and CAS. However, the decrease in T2 is measurable prior to the decrease in CAS (when this was still 7) and only later reduced to 2. As discussed earlier, this is because of the lack of sensitivity of the CAS to change—there is only a reduction in the score when signs or symptoms are entirely resolved, not when there is a substantial improvement with residual symptoms.

This study demonstrates the feasibility of measuring FF in EOMs, thereby quantifying the relative contribution of fat and water to the T2 signal. The increased FF within EOMs in TED patients is consistent with the suggestion that there is fatty infiltration and/or adipogenesis within the muscles in the course of the disease [21, 22]. Interestingly, the data also demonstrate that EOMs in healthy volunteers contain a measurable proportion of fat. Ultimately, higher FFs may reflect chronicity of the disease. Thus, the ability to differentiate between the contribution played by water and hyaluronic acid (presumed to equate to inflammation) and fat (presumed to equate to chronicity) may help guide decisions regarding treatment.

The limitations of this approach include the manual placement of regions of interest within a single slice through each EOM. Superior and lateral rectus muscles have smaller cross-sectional areas, the superior rectus muscle is not easily separated from levator palpebrae superioris and lateral rectus lies at an oblique angle to the coronal plane in which images are acquired. The regions of interest could also include voxels at the edges of muscle, where there may be artefact related to the interface between muscle and orbital fat. These factors may account for some of the variability in measurements of FF in both patients and controls. However, the variability is also consistent with the asymmetric and heterogeneous nature of the disease. There may also be an unknown effect of age as the control population are not age matched. Furthermore, the patient population was unselected in terms of clinical severity and chronicity.

In conclusion, quantitative T2-relaxation time measurements of EOMs in TED correlate with CAS and MRI SIR measurements. However, T2-relaxation measurements better discriminate abnormal from normal populations. Although this is an attractive objective measure of disease activity, it may be confounded by the presence of fat within muscles. We address this by concurrent measurement of FFs such that this combined MR approach may allow for improved quantification of disease activity and for monitoring within individual patients for treatment effect. Our physiological assessments of inflammation occasionally contradicted CAS scores, raising questions about the validity of clinical scoring of disease activity.

Acknowledgements We would like to acknowledge Dr. Paul Meyer, Dr. Nagui Antoun, and the MRI Radiographers.

Funding The project was supported by the Addenbrooke's Charitable Trust and the National Institute for Health Research [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust]. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Author contributions The team conceived of the idea for this study and all authors have contributed to the data gathering, analysis and writing.

Compliance with ethical standards

Conflict of interest All authors are doctors who assess patients with thyroid eye disease either clinically or radiologically. There are no other conflict of competing interests and a unified declaration form is available on request.

Ethics statement This study was approved by the local Research Ethics Committee.

License The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors an exclusive licence on a worldwide basis to permit this article to be published.

Guarantor Dr. Das serves as guarantor of this work. It is an honest, accurate and transparent account of the study being reported; no important aspects of the study have been omitted.

References

- Pitz S, Muller-Forell W. Orbital imaging. In: Wiersinga WM, Kahaly GJ, editors. Graves' orbitopathy: a multidisciplinary approach—questions and answers. Basel, Karger; 2017. p 61–73.
- Konuk O, Anagnostis P. Diagnosis and differential of Graves' orbitopathy. In: Wiersinga WM, Kahaly GJ, editors. Graves' orbitopathy: a multidisciplinary approach—questions and answers. Basel, Karger; 2017. p 74–92.
- Mourits MP, Koornneef L, Wiersinga WM, Prummel MF, Berghout A, van der Gaag R. Clinical criteria for the assessment of disease activity in Graves' ophthalmopathy: a novel approach. *Br J Ophthalmol*. 1989;73:639–44.
- Gorman CA. The measuring of changes in Graves' ophthalmopathy. *Thyroid*. 1998;8:539–43.
- Dickinson AJ, Hintschich C. Clinical manifestations. In: Wiersinga WM, Kahaly GJ, editors. Graves' orbitopathy: a multidisciplinary approach—questions and answers. Basel, Karger; 2017. p 1–25.
- Dickinson AJ, Perros P. Controversies in the clinical evaluation of active thyroid-associated orbitopathy: use of a detailed protocol with comparative photographs for objective assessment. *Clin Endocrinol (Oxf)*. 2001;55:283–303.
- Dolman PJ. Grading severity and activity in thyroid eye disease. *Ophthalmic Plast Reconstr Surg*. 2018;34 (4S Suppl 1):S34–40.
- Werner SC. Modification of the classification of the eye changes of Graves' disease: recommendations of the Ad Hoc Committee of the American Thyroid Association. *J Clin Endocrinol Metab*. 1977;44:203–4.

9. Smith TJ, Kahaly GJ, Ezra DG, Fleming JC, Dailey RA, Tang RA, et al. Teprotumumab for thyroid-associated ophthalmopathy. *N Engl J Med*. 2017;376:1748–61.
10. Roos JC, Murthy R. Personal communication. 2018.
11. Roos JCP, Paulpandian V, Murthy R. Serial TSH-receptor antibody levels to guide the management of thyroid eye disease: the impact of smoking, immunosuppression, radio-iodine, and thyroidectomy. *Eye (Lond)*. 2018 Nov 6. <https://doi.org/10.1038/s41433-018-0242-9>. [Epub ahead of print].
12. Tortora F, Prudente M, Cirillo M, Elefante A, Belfiore MP, Romano F, et al. Diagnostic accuracy of short-time inversion recovery sequence in Graves' ophthalmopathy before and after prednisone treatment. *Neuroradiology*. 2014;56:353–61.
13. Pajak M, Loba P, Wieczorek-Pastusiak J, Antosik-Biernaka A, Stefańczyk L, Majos A. Signal intensity and T2 time of extraocular muscles in assessment of their physiological status in MR imaging in healthy subjects. *Pol J Radiol*. 2012;77:7.
14. Mayer E, Herdman G, Burnett C, Kabala J, Goddard P, Potts MJ. Serial STIR magnetic resonance imaging correlates with clinical score of activity in thyroid eye disease. *Eye*. 2001;15:313–8.
15. Mancini L, Rajendram R, Uddin J, Lee RW, Rose GE, Youstry T, et al. Extra-orbital muscle T2 relaxation time and clinical activity in thyroid eye disease. In: Paper presented at International Society for Magnetic Resonance in Medicine. https://cds.ismrm.org/protected/10MProceedings/files/876_3843.pdf Accessed 28 Oct 2018.
16. Ohnishi T, Noguchi S, Murakami N, Tajiri J, Harao M, Kawamoto H, et al. Extraocular muscles in Graves ophthalmopathy: usefulness of T2 relaxation time measurements. *Radiology*. 1994;190:857–62.
17. Reeder SB, Sirlin CB. Quantification of liver fat with magnetic resonance imaging. *Magn Reson Imaging Clin N Am*. 2010;18:337–57–ix.
18. Hussain HK, Chenevert TL, Londy FJ, Gulani V, Swanson SD, McKenna BJ, et al. Hepatic fat fraction: MR imaging for quantitative measurement and display—early experience. *Radiology*. 2005;237:1048–55.
19. Politi LS, Godi C, Cammarata G, Ambrosi A, Iadanza A, Lanzi R, et al. Magnetic resonance imaging with diffusion-weighted imaging in the evaluation of thyroid-associated orbitopathy: getting below the tip of the iceberg. *Eur Radiol*. 2014;24:1118–26.
20. Mayer EJ, Fox DL, Herdman G, Hsuan J, Kabala J, Goddard P, et al. Signal intensity, clinical activity and cross-sectional areas on MRI scans in thyroid eye disease. *Eur J Radiol*. 2005;56:20–4.
21. Li H, Fitchett C, Kozdon K, Jayaram H, Rose GE, Bailly M, et al. Independent adipogenic and contractile properties of fibroblasts in Graves' orbitopathy: an in vitro model for the evaluation of treatments. *PLoS ONE* 2014;9:e95586.
22. Lee BW, Kumar VB, Biswas P, Ko AC, Alameddine RM, Granet DB, et al. Transcriptome analysis of orbital adipose tissue in active thyroid eye disease using next generation RNA sequencing technology. *Open Ophthalmol J*. 2018;12:41–52.
23. Polito E, Leccisotti A. MRI in Graves orbitopathy: recognition of enlarged muscles and prediction of steroid response. *Ophthalmologica*. 1995;209:182–6.
24. Ozgen A, Alp MN, Ariyürek M, Tütüncü NB, Can I, Günalp I. Quantitative CT of the orbit in Graves' disease. *Br J Radiol*. 1999;72:757–62.
25. Bijlsma WR, Mourits MP. Radiologic measurement of extraocular muscle volumes in patients with Graves' orbitopathy: a review and guideline. *Orbit*. 2006;25:83–91.
26. Alp MN, Ozgen A, Can I, Cakar P, Gunalp I. Colour Doppler imaging of the orbital vasculature in Graves' disease with computed tomographic correlation. *Br J Ophthalmol*. 2000;84:1027–30.
27. Kuo PH, Monchamp T, Deol P. Imaging of inflammation in Graves' ophthalmopathy by positron emission tomography/computed tomography. *Thyroid*. 2006;16:419–20.
28. Di Maria C, Allen J, Dickinson J, Neoh C, Perros P. Novel thermal imaging analysis technique for detecting inflammation in thyroid eye disease. *J Clin Endocrinol Metab*. 2014;99:4600–6.
29. Tortora F, Cirillo M, Ferrara M, Belfiore MP, Carella C, Caranci F, et al. Disease activity in Graves' ophthalmopathy: diagnosis with orbital MR imaging and correlation with clinical score. *Neuroradiol J*. 2013;26:555–64.
30. Lingam RK, Mundada P, Lee V. Novel use of non-echo-planar diffusion weighted MRI in monitoring disease activity and treatment response in active Grave's orbitopathy: an initial observational cohort study. *Orbit*. 2018;37:325–30.
31. Laitt RD, Hoh B, Wakeley C, Kabala J, Harrad R, Potts M, et al. The value of the short tau inversion recovery sequence in magnetic resonance imaging of thyroid eye disease. *Br J Radiol*. 1994;67:244–7.
32. Kaichi Y, Tanitame K, Itakura H, Ohno H, Yoneda M, Takahashi Y, et al. Orbital fat volumetry and water fraction measurements using t2-weighted FSE-IDEAL imaging in patients with thyroid-associated orbitopathy. *Am J Neuroradiol*. 2016;37:2123–8.
33. Taoka T, Sakamoto M, Nakagawa H, Fukusumi A, Iwasaki S, Taoka K, et al. Evaluation of extraocular muscles using dynamic contrast enhanced MRI in patients with chronic thyroid orbitopathy. *J Comput Assist Tomogr*. 2005;29:115–20.
34. Sun B, Song L, Wang X, Li J, Xian J, Wang F, et al. Lymphoma and inflammation in the orbit: diagnostic performance with diffusion-weighted imaging and dynamic contrast-enhanced MRI. *J Magn Reson Imaging*. 2017;45:1438–45.
35. Jiang H, Wang Z, Xian J, Li J, Chen Q, Ai L. Evaluation of rectus extraocular muscles using dynamic contrast-enhanced MR imaging in patients with Graves' ophthalmopathy for assessment of disease activity. *Acta Radiol*. 2012;53:87–94.
36. Guo Y, Huo L, Wang P, Huang L, Chai C, Sun F, et al. Evaluating the microcirculation of normal extraocular muscles using quantitative dynamic contrast-enhanced magnetic resonance imaging. *J Comput Assist Tomogr*. 2016;40:419–23.