

Non-EPI-DWI for Detection, Disease Monitoring, and Clinical Decision-Making in Thyroid Eye Disease

C. Feeney, R.K. Lingam, V. Lee, F. Rahman, and S. Nagendran



ABSTRACT

BACKGROUND AND PURPOSE: The Clinical Activity Score is widely used to grade activity of thyroid eye disease and guide treatment decisions, but as a subjective measurement and being confined to the anterior orbit, it has limitations. Non-EPI-DWI of the extraocular muscles may offer advantages as a functional imaging technique with reduced skull base artifacts, but the correlation with the Clinical Activity Score and patient outcome is unknown. Our aim was to establish the correlation between the Clinical Activity Score and non-EPI-DWI and to describe the additional value provided by adjunctive non-EPI-DWI in making clinical decisions.

MATERIALS AND METHODS: This was a retrospective longitudinal study of 31 patients seen in a multidisciplinary thyroid eye disease clinic during 5 years who had at least 1 ophthalmic and endocrine assessment including the Clinical Activity Score and a non-EPI-DWI ADC calculation. The Spearman rank correlation coefficient was used to determine the relationship between the Clinical Activity Score and non-EPI-DWI. A patient flow chart was constructed to evaluate clinical decision-making, and receiver operating characteristics were generated.

RESULTS: From 60 non-EPI-DWI scans, 368 extraocular muscles were selected for analysis. There was a significant positive correlation between the Clinical Activity Score and ADC ($r_s = 0.403$; 95% CI, 0.312–0.489; $P < .001$). ADC values were significantly higher in the Clinical Activity Score ≥ 3 group compared with the Clinical Activity Score < 3 group ($P < .001$). Our patient flow chart identified a third intermediate-severity cohort in which the non-EPI-DWI was particularly useful in guiding clinical decisions.

CONCLUSIONS: The non-EPI-DWI correlated well with the Clinical Activity Score in our patients and was a useful adjunct to the Clinical Activity Score in making clinical decisions, especially in patients with intermediate activity and severity of thyroid eye disease.

ABBREVIATIONS: CAS = Clinical Activity Score; DON = dysthyroid optic neuropathy; EOM = extraocular muscle; ROC = receiver operating characteristic; TED = thyroid eye disease

Thyroid eye disease (TED) is one of the most common inflammatory diseases of the orbit affecting approximately 25% of individuals presenting with systemic Graves disease.^{1,2} TED can be a sight-threatening disease with profound physical and psychological consequences.³ The time course of the disease typically follows an initial active phase, potentially amenable to immunosuppressive

treatment, and a later inactive phase in which residual structural orbital disease is usually surgically managed.⁴

Disease activity is often guided by the Clinical Activity Score (CAS) based on scoring of mainly subjective markers of disease, with a score of ≥ 3 often denoting cases in which systemic immunosuppression should be considered.^{5,6} Despite its wide use, the CAS has limitations. The CAS can underestimate activity in the posterior orbit, and most important, sight-threatening dysthyroid optic neuropathy (DON) may occur in the presence of a low CAS score.⁷

Advances and availability in MR imaging have resulted in a trend toward the application of MR imaging to TED to inform disease activity alongside CAS.^{8,9} Various MR imaging modalities have been applied in this setting and include T2-weighted extraocular muscle (EOM) signal intensities, dynamic contrast-enhanced MR imaging, and fat-suppression techniques such as the STIR sequence.¹⁰⁻¹³ Results from these reports have been variable, and some techniques have additional limitations that

Received October 13, 2019; accepted after revision May 16, 2020.

From the Eye Department (C.F., V.L., F.R., S.N.), Central Middlesex Hospital, London, UK; Imperial Centre for Endocrinology (C.F.), North West Thames, UK; and Department of Radiology (R.K.L.), Northwick Park & Central Middlesex Hospital, London Northwest University Healthcare National Health Service Trust, London, UK.

C.F. was funded by a Medical Research Council research grant during the period of data collection.

Please address correspondence to Vickie Lee, MD, Eye Department, Central Middlesex Hospital, London, Acton Lane, Park Royal, London, NW10 7NS, UK; e-mail: v.lee@imperial.ac.uk; vickie.lee@nhs.net

Indicates open access to non-subscribers at www.ajnr.org

<http://dx.doi.org/10.3174/ajnr.A6664>

include the adverse events associated with contrast administration and necessary standardization corrections to allow reliable comparison between scans (eg, normalization of signal intensities with those of the nearby temporalis muscle).^{14,15}

DWI is an MR imaging technique that measures movement of water molecules within tissues to calculate an ADC, in which higher ADC values indicate greater facilitated diffusion reflective of an underlying active inflammatory process.¹⁴ Echo-planar DWI has shown promise in the evaluation of active TED, and studies have demonstrated higher ADC values in the EOMs of patients with TED compared with healthy controls.^{14,16,17} In studies comparing active with inactive TED, the ADC is reported to be higher in active disease and is elevated at an early stage in the disease before activity is detected on routine MR imaging, suggesting the potential utility of ADC evaluation at an early time point in the disease trajectory.^{16,17}

Non-EPI-DWI, however, has the potential of higher resolution images and a reduction in air-bone interface artifact distortion.¹⁸ These advantages potentially allow greater clarity of orbital images and more reliable measurement of ADC values, and non-EPI-DWI has replaced its EPI counterpart in assessing other skull base pathologies such as middle ear cholesteatoma.^{19,20} In TED, we have demonstrated in a case series that non-EPI-DWI was a useful adjunct in our management of selected cases of moderate-to-severe and active TED.²¹ Acknowledging that conventional MR imaging STIR signal intensity ratios are a commonly used MR imaging technique in TED, we previously conducted a preliminary study to establish that signal intensity ratios and ADC are positively correlated and that there was good interobserver correlation within each measurement in our center.^{21,22}

The objectives of this study were to build on our previous work to do the following: 1) investigate the correlation of non-EPI-DWI with CAS; 2) describe the clinical decisions made on the basis of non-EPI-DWI data in a tertiary referral multidisciplinary setting; and 3) explore the diagnostic potential of non-EPI-DWI in mild and sight-threatening disease.

MATERIALS AND METHODS

Subjects

This retrospective observational study examined all adult patients who were referred to a multidisciplinary TED clinic from inception of the service in October 2012 to data base lock in April 2017 and fulfilled the following inclusion criteria: 1) suspected diagnosis of TED in combination with either signs or symptoms of proptosis, ocular motility disturbance, reduced visual acuity, and/or known thyroid disease; 2) at least 1 endocrine and ophthalmologic assessment and at least 1 CAS measurement; 3) at least 1 non-EPI-DWI orbital MR imaging scan with at least 1 measurement of an EOM ADC value. This study design was reviewed and approved by the institutional research and development review board (Northwest University Healthcare National Health Service Trust). The requirement for informed consent was waived.

Clinical Assessment

Patients were evaluated in a multidisciplinary TED clinic with a consultant ophthalmic and oculoplastic specialist with an interest in TED (V.L.), an endocrinologist (C.F.), and orthoptist. The

CAS score is a validated score to grade disease activity in TED.⁵ CAS assessments were performed by the same observer (V.L. with >20 years' experience) at each visit and scores of ≥ 3 were considered moderate-to-severe and active cases of TED.⁶ In the case of multiple CAS measurements, the score nearest in date to MR imaging was selected for analysis (mean time from CAS to MR imaging, 13.2 ± 21.9 days).

All patients had at least 1 endocrine assessment with optimization of the thyroid status with either antithyroid medication (carbimazole first-line and propylthiouracil second-line) or thyroxine replacement or both (ie, block and replace) in patients with moderate-to-severe and active eye disease or in patients with unstable thyroid function.

The following clinical data were collected retrospectively for each patient: 1) date of birth; 2) sex; 3) ethnicity; 4) medical history including a history of thyroid disorder and previous radioiodine treatment; 5) a family history of Graves disease; 6) smoking status; 7) thyroid status including antibody assessment; 8) medications at initial visit; 9) CAS measurements closest in time to DWI scans; 10) formal DWI radiology report; and 11) TED treatments between scans (eg, intravenous methylprednisolone, orbital radiation therapy, second-line immunosuppressive treatment, and surgical interventions).

Imaging Technique

MR imaging was performed on a 1.5T superconductive unit (Magnetom Avanto; Siemens) using a standard Head Matrix coil. In all patients, a 3-mm-thick HASTE DWI sequence was acquired in the coronal plane (TR = 900 ms; TE = 118 ms; matrix = 192×86 ; FOV = 145 mm; 18 averages; EPI spacing = 6.28 ms; bandwidth = 465 Hz/pixel; b factors, 0 and 1000 s/mm²). Coronal 3-mm-thick STIR images (TR = 5640 ms; TE = 91 ms; matrix = 256×112 ; FOV = 145 mm) were obtained. The STIR, $b = 0$, and $b = 1000$ diffusion-weighted images were copy-referenced to ensure the same section position to allow optimal image evaluation and measurement. Following acquisition, an ADC map was calculated using the diffusion scan raw data.

Imaging Evaluation

A specialist head and neck radiologist, blinded to the CAS score (R.K.L. with 14 years' experience), examined all MR images before selecting the non-EPI-DWI section that best illustrated EOMs with higher ADC on visual survey, corroborated by an observer (C.F.) also blinded to CAS score. We have previously observed good interobserver agreement for non-EPI-DWI ADC in 23 actively inflamed muscles in cases of TED in our center (intraclass coefficient = 0.97; 95% CI, 0.94–0.98).²² Freehand ROIs were contoured within the inner border of all visualized active EOMs to determine the ADC (\pm SD) value. In addition, ADC values were derived for the medial and inferior recti bilaterally on all scans, irrespective of the initial survey, because TED appears to have a predilection for these EOMs.¹⁴ The oblique muscles were excluded from analysis due to concerns about the reliability of measurement in an oblique plane as well as being atypical sites for TED involvement. Right and left values in each individual were analyzed separately to reflect that disease can be unilateral and asymmetric. Therefore, the minimum number of EOM measurements per scan was 4 (ie, right and left

inferior and medial recti alone) or a maximum of 8 (ie, inferior and medial recti as well as the superior and/or lateral recti bilaterally). Each scan was assessed in an identical fashion regardless of whether it was a baseline or follow-up scan.

Subgroup Analyses

Three clinically defined subgroups were selected for further analysis. Cohort 1 (96 EOMs, $n = 12$) was defined as mild/possibly active TED based on clinical criteria, a CAS of ≤ 1 , and no more than 1 scan. Cohort 2 (100 EOMs, $n = 8$) referred to individuals who were considered to have definite TED but fell within a mild-to-moderate and active category of severity based on CAS 1–3. These subjects had an initial scan at presentation and at least 1 follow-up scan. The remaining cohort 3 (172 EOMs, $n = 11$) had moderate-to-severe and active TED with CAS scores of ≥ 3 . These subjects also had at least 2 scans. For selected analyses, a DON subgroup ($n = 5$) was defined as those patients with acuity loss and requiring high-dose pulse methylprednisolone and/or emergency orbital decompression.

Patient Flow Chart

To describe the clinical decisions made on the basis of clinical assessment, the CAS, and the non-EPI-DWI result (determined globally as active or inactive based on the initial radiology report by R.K.L.), we constructed a flow chart to follow the patient journey for each of cohorts 1–3 and documented the patient outcome at study close. This was performed to determine whether non-EPI-DWI had any observed advantage for patients beyond the CAS and clinical assessment alone.

Receiver Operating Characteristic

To evaluate the potential utility of ADC as a diagnostic test for DON or to exclude mild/possible TED, we generated receiver operating characteristic (ROC) curves for the 2 following scenarios: 1) cases of DON versus no DON, and 2) cohort 1 (mild/possible TED) versus cohort 3 (moderate-severe TED).

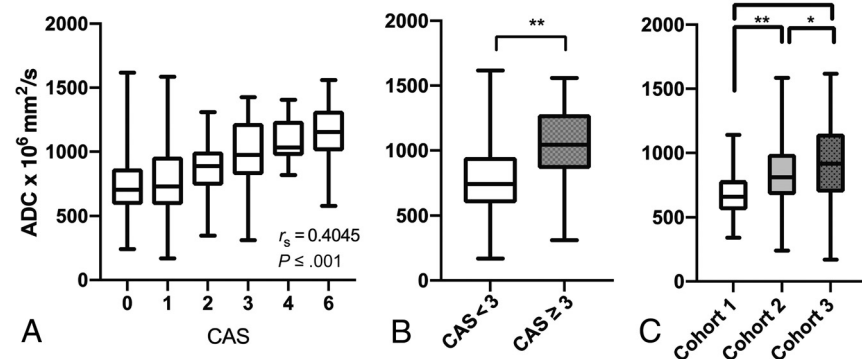


FIG 1. Correlation between CAS and ADC and clinically meaningful group-wise comparisons of ADC and CAS in the overall cohort. Box plots demonstrate a positive Spearman rank correlation coefficient between CAS and ADC of individual EOMs ($n = 368$) (A) and significantly greater ADC values in CAS ≥ 3 ($n = 70$) (B) compared with CAS < 3 ($n = 298$) groups. Two asterisks indicate $P < .001$, 2-tailed. C, Significantly greater ADC values of all EOMs in cohort 3 compared with those in cohort 1 ($P < .001$) and cohort 2 ($P = .03$). For cohort 2, ADC values were significantly higher than those in cohort 1 ($P < .001$) but lower than those in cohort 3 ($P = .03$). Single asterisk indicates $P < .05$.

Statistical Analysis

Statistical analyses, including ROC generation, were performed using GraphPad Prism software, Version 8.0.0 (GraphPad Software). Demographic data are presented as number and percentage; group averages are reported as mean \pm SD, range, and Fisher exact/Mann-Whitney U test or 1-way ANOVA to compare groups. To calculate correlations of CAS and ADC values, we used a paired Spearman rank correlation coefficient statistical test (r_s , confidence interval, and P value). For subgroup comparisons of non-normally distributed data, we applied the Mann-Whitney U test. P values $< .05$ were considered significant.

RESULTS

Patient Characteristics

Of the 88 patients seen in the multidisciplinary clinic, 31 met the inclusion criteria (60 DWI scans, 368 EOMs) and were included in the final analysis (mean age, 48.2 years; range, 20.2–79.8 years; 22/31 [70.9%] women; 32% smokers). Given that only a subset of the total cohort warranted specialized imaging to inform clinical care, most exclusions (55/57, 96%) were due to no DWI scan being available for review, with the remainder (2/57, 4%) due to no documented CAS.

There was strong representation from ethnicities that were not white, in keeping with demographics of a large metropolis and included 32% Black Caribbean, 32% white (UK and Europe), 23% Indian/Asian, and 10% Middle Eastern. At presentation to the clinic, 20/31 (64.5%) were on treatment for confirmed Graves thyrotoxicosis (16/20 carbimazole and 4/20 prophyllthiouracil), 3/31 (9.7%) were on levothyroxine, and 8/31 (25.8%) had normal thyroid function and were not receiving treatment.

There was a significant positive correlation of moderate effect size between CAS and ADC across all 368 EOMs ($r_s = 0.403$; 95% CI, 0.312–0.489; $P < .001$) (Fig 1A). Given that CAS ≥ 3 often denotes a cutoff for clinically active moderate-to-severe and active TED, a group-wise comparison was made between CAS < 3 (EOMs, $n = 298$) and CAS ≥ 3 (EOMs, $n = 70$) groups. ADC values were significantly higher in the CAS ≥ 3 group (mean ADC, 913 ± 312) compared with the CAS < 3 group (mean ADC, 760 ± 239) ($P \leq .001$) (Fig 1B).

Subgroup Analyses

Demographic data are shown in Table 1. There was greater representation from Black Caribbean racial groups in cohorts 2 and 3 versus cohort 1, but this did not reach statistical significance. All patients in cohort 3 had a positive thyroid auto-antibody result and were taking antithyroid medication at presentation.

ADC values in cohort 3 were significantly higher than those in cohorts 1 (96 EOMs, $P < .001$) and 2 (100 EOMs, $P = .03$). In cohort 2, ADC values were significantly higher than

Table 1: Patient characteristics (cohorts 1–3)

	No.	Cohort 1	No.	Cohort 2	No.	Cohort 3
Age at initial scan (yrs)	12	43.9	8	47.0	11	53.7
SD + range		15.4 (20.2–74.4)		12.9 (33.9–71.9)		20.8 (22.6–79.8)
n% female	12	9 (75.0%)	8	5 (62.5%)	11	8 (72.7%)
n% Afro-Caribbean	11 ^a	1 (9.1%)	8	3 (37.5%)	11	5 (45.5%)
n% current smokers	12	4 (33.3%)	8	3 (37.5%)	11	3 (27.3%)
n% positive family history	12	4 (33.3%)	8	3 (37.5%)	11	1 (9.1%)
n% positive autoantibody ^b	9 ^c	5 (55.5%)	8	4 (50.0%)	10 ^d	10 (100.0%)
n% antithyroid medication ^e	12	6 (50.0%)	8	6 (75.0%)	11	11 (100.0%)
n% euthyroid ^f	12	6 (50.0%)	8	2 (25.0%)	11	0 (0.0%)
n% DON	12	0 (0.0%)	8	1 (12.5%)	11	4 (36.3%)
n% IV methylprednisolone	12	0 (0.0%)	8	2 (25.0%)	11	11 (100.0%)
n% second-line immunosuppression ^g	12	0 (0.0%)	8	0 (0.0%)	11	3 (27.3%)
n% orbital radiotherapy	12	0 (0.0%)	8	0 (0.0%)	11	7 (63.6%)
No. of scans	12	1	8	2.5	11	2.5
SD + range		0 (1.0–1.0)		0.9 (2.0–4.0)		0.8 (2.0–4.0)
Initial CAS	12	0.5	8	1.6	11	3.6
SD + range		0.5 (0.0–1.0)		0.5 (1.0–2.0)		1.1 (3.0–6.0)
CAS: 1st Follow-up	NA	NA	8	0.5	11	2.1
SD + range				0.8 (0–2)		1.9 (0–6)
Baseline ADC all EOMs	96	678	54	811	65	691
SD + range		171 (340–1141)		256 (311–1426)		208 (240–1088)
ADC all EOMs:1st Follow-up	NA	NA	52	770	63	873
SD + range				236 (340–1321)		319 (169–1585)
Time between 1st & 2nd scan (Months)	NA	NA	8	10.5	11	9.5
SD + range				3.1 (5.5–16.2)		7.8 (2.0–29.0)
Final CAS	12	NA	8	0.4	11	1.6
SD + range				0.1 (0.0–1.0)		2.0 (0.0–6.0)
Total follow-up period (Months)	12	NA	8	39.3	11	48
				27.7 (16.3–97.4)		49.3 (8.3–163.0)

Note:—NA indicates not applicable.

^a Data unrecorded ($n = 1$).

^b Either thyroxine peroxidase or thyroxine-stimulating hormone receptor antibody.

^c Data missing ($n = 3$).

^d Data missing ($n = 1$).

^e Either carbimazole, propylthiouracil, or both.

^f Normal thyroid function and no history of thyroid abnormalities.

^g Mycophenolate or hydroxychloroquine. Cohort 1 = mild/possibly active (CAS 0 or 1), cohort 2 = mild-to-moderate and active (CAS 1–3), cohort 3 = moderate-to-severe and active (CAS ≥ 3).

those in cohort 1 ($P < .001$) and lower than those in cohort 3 (172 EOMs, $P = .03$) (Fig 1C).

Subgroup Results: Clinical Course

There were differences in the clinical courses of patients in each cohort. In cohort 1, there were no patients who developed moderate-to-severe and active TED requiring medical or surgical treatment, and to our knowledge, no patients re-presented with symptoms during a 6- to 18-month follow-up.

In cohort 3, all patients had at least 1 course of IV methylprednisolone compared with 2/8 (25%) patients in cohort 2 ($P = .001$). Seven of 11 patients in cohort 3 received orbital radiation therapy compared with 0/8 patients in cohort 2 ($P = .004$). Greater numbers of patients in cohort 3 (3/11, 27.3%) required second-line immunosuppressants (mostly mycophenolate) compared with none in cohort 2. Four patients (36.3%) developed DON in cohort 3 and 1/8 (12.5%) in cohort 2, all necessitating high-dose pulsed methylprednisolone with 2/5 requiring surgical decompression (Table 1).

Patient Flow Chart

Results of the patient flow chart are shown in Fig 2. In cohort 1, most patients, 10/12 (83.3%), showed concordance between a

mild clinical picture and an inactive scan and were discharged after a follow-up clinical review. The remaining 2 showed discordance (ie, an active scan) and were followed more closely but remained stable.

Non-EPI-DWI appeared to have the most value as an adjunct to the CAS in determining treatment and monitoring decisions in cohort 2 (ie, patients with active disease but not meeting the CAS criteria for systemic immunosuppression). Those patients in cohort 2 who also had an active scan (ie, 7/8, 87.5%) were not treated with immunosuppressive therapies, as were those with CAS < 3 , but a more conservative approach of clinical surveillance in combination with a second interval scan at 3–6 months was pursued. Of these, 2 patients remained clinically and radiologically active at the next review to justify systemic immunosuppression despite CAS < 3 . One of these patients subsequently developed DON, and the other required further courses of immunosuppression to control the disease (latter patient shown in Fig 3B).

Only 1/11 (9.9%) patients in cohort 3 had a discordant clinical and radiologic picture (ie, CAS > 3 but scan inactive). This patient was treated regardless with IV methylprednisolone on the basis of clinical judgment (CAS = 4); however, the CAS remained high during the 18-month follow-up period, raising the possibility of a type I error in the CAS. All other patients in this group had an

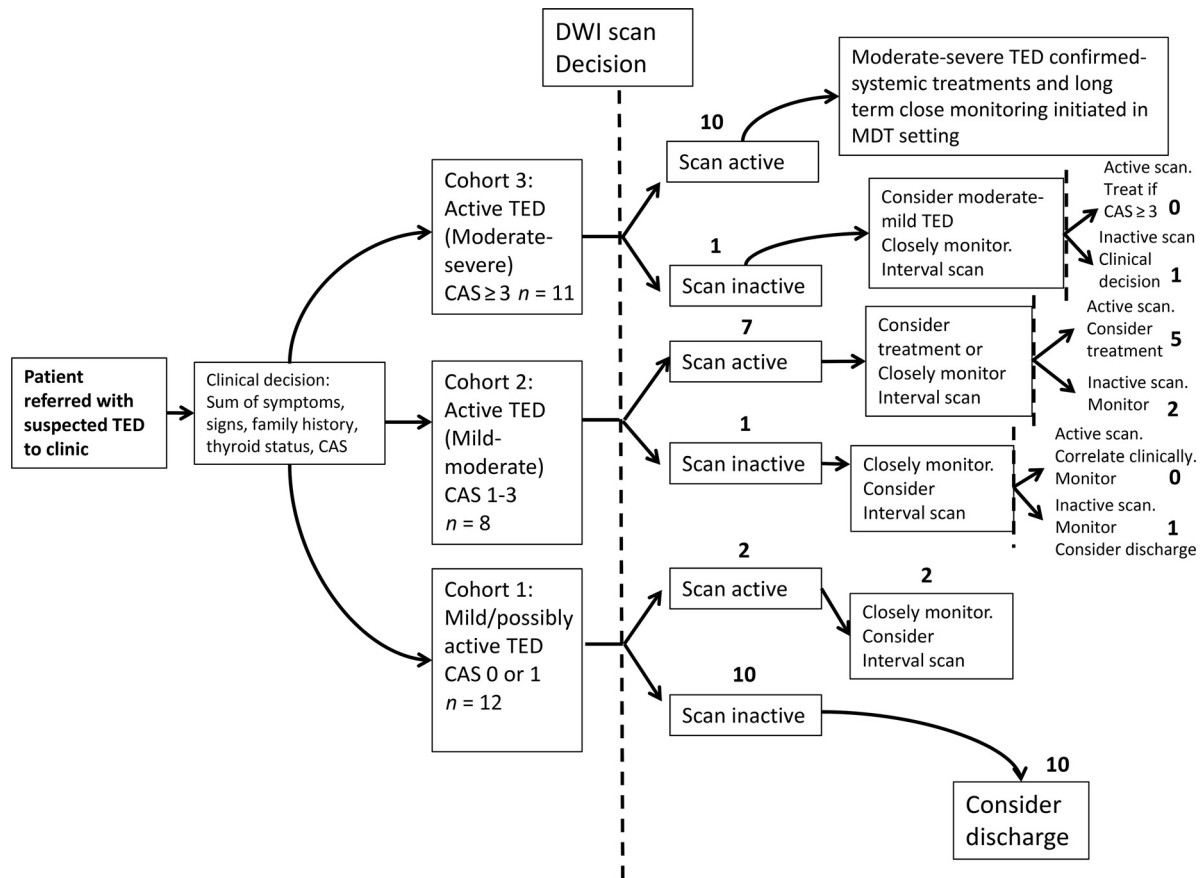


FIG 2. Patient flow chart for clinical decisions influenced by the reported activity of non-EPI-DWI scans in conjunction with CAS and clinical assessment. Whole numbers denote the number of patients. The *dashed line* represents the opportunity for a non-EPI-DWI-informed clinical decision. MDT indicates multidisciplinary team.

active scan and received systemic treatment, radiation therapy, and surgical intervention as described earlier. All of these patients had a protracted clinical course, and 3/10 (30%) relapsed at some point during the follow-up period. A representative non-EPI-DWI from a patient within this cohort (who was treated for DON) is shown in Fig 3C.

ROC

ADC performed reasonably well in differentiating cohort 1 from cohort 3, with an estimated ADC cutoff of <833 to give an 83% sensitivity and 61% specificity in diagnosing mild/possibly active disease (Fig 4). An optimum ADC cutoff to diagnose DON was estimated at >1154 to give 100% sensitivity and 96% specificity; however, the analysis was severely limited by the low number of DON cases included in the analysis ($n = 5$).

DISCUSSION

Our results build on our previous work to demonstrate that a significant positive correlation exists among the CAS, the commonly used the TED activity score, and ADC derived from non-EPI-DWI of EOMs in cases of suspected active TED.^{21,22} Previous studies have also demonstrated this finding using echo-planar-DWI, but to our knowledge, this is the first demonstrating a positive relationship, using non-EPI-DWI.^{16,17}

Non-EPI-DWI has potential advantages over echo-planar-DWI, which include greater resolution, thinner-section images,

and refocusing pulses for every measured EPI-DWI sequence that reduces air-bone interface artifacts and distortion.^{19,23-25} Given DWI scanning parameters (including the number of b-values used) are standardized, there is no requirement for normalization of ADC values with those of the adjacent temporalis muscles, unlike other MR imaging modalities, and DWI obviates contrast administration.

Using non-EPI-DWI in our clinic, we identified 3 broad groups of patients with distinct characteristics, clinical courses, CAS scores, and imaging features. Although the moderate-to-severe and active group is a well-recognized TED group often defined as $CAS \geq 3$ and usually requiring systemic immunosuppressive therapies, our findings have highlighted, for the first time, 2 further subgroups with $CAS < 3$ in which clinical decision-making can be more challenging. While we made decisions primarily on the basis of clinical judgment and in line with European guidance,⁶ we found, from our own experience, that agreement or disagreement with non-EPI-DWI allowed us to make more confident decisions regarding discharge, monitoring, and treating with immunosuppressive therapy with $CAS < 3$. However, a larger follow-up study is needed to determine what impact the addition of non-EPI-DWI to a management algorithm has on long-term patient outcomes.

We believe that non-EPI-DWI was a useful adjunct for our patients, in part because the CAS can sometimes over- or underestimate true disease activity, even in very experienced hands and

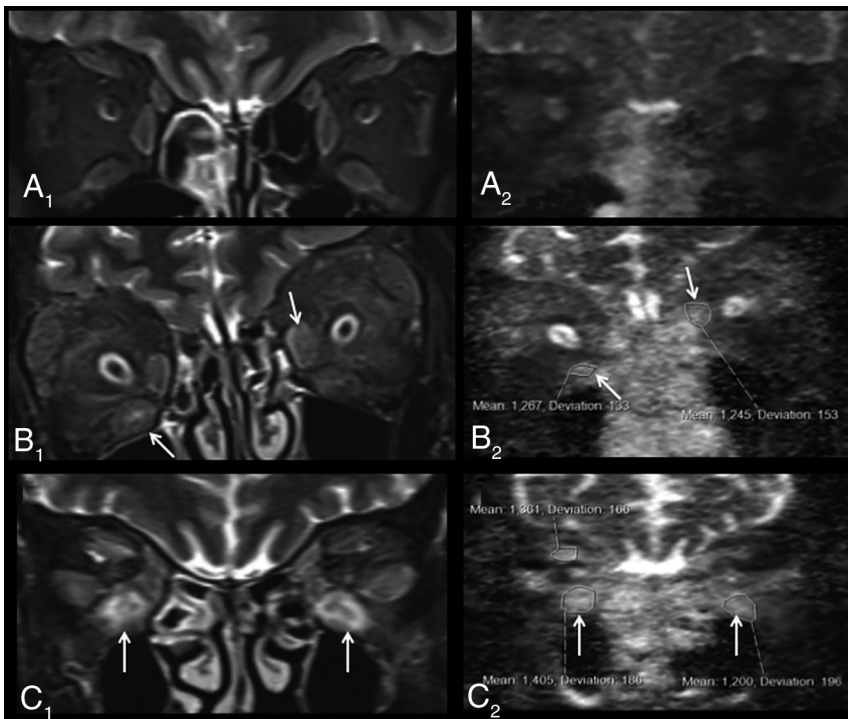


FIG 3. Representative examples of non-EPI-DWI of orbital EOMs in patients from each of cohorts 1, 2, and 3 alongside STIR MR imaging. A, Coronal orbital/EOM MR imaging STIR image (A_1) and a non-EPI-DWI ADC image (A_2) show mild enlargement of the extraocular muscles with mild increased signal and ADC values (right inferior rectus muscle = 590; left inferior rectus muscle = 540), labeled an inactive scan (cohort 1). STIR image (B_1) and ADC image (B_2) show moderate-to-marked enlargement of the extraocular muscles and increased signal and ADC values, notably at the right inferior and left medial recti muscles (arrows) (cohort 2). STIR image (C_1) and ADC image (C_2) show moderate-to-marked enlargement of the extraocular muscles and markedly increased signal and ADC values notably at the inferior recti muscles (arrows) (cohort 3).

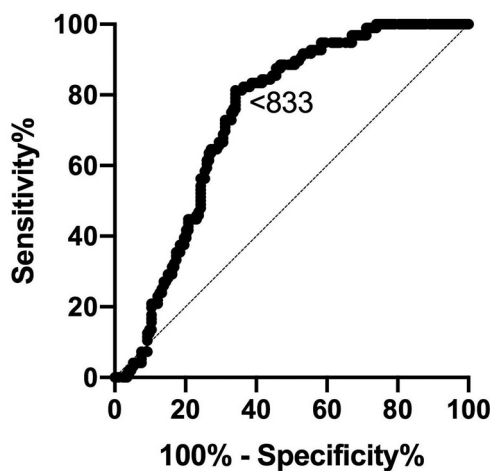


FIG 4. ROCs for ADC values obtained in subjects with no/possible disease (cohort 1) and moderate-severe disease (cohort 3) (area under the curve = 0.737; 95% CI, 0.68–0.80; $P < .001$). Diagonal line represents a line of no discrimination between disease states. Arrowed number represents the ADC value.

when controlling for interrater variability in clinical judgment as we did in this study. For example, one of our patients who presented with $CAS < 3$ was not treated immediately with systemic

immunosuppression but eventually required treatment after active progression was noted on non-EPI-DWI. This patient subsequently developed and was successfully treated for DON. In this case, the CAS, being largely restricted to signs in the anterior orbit, may have underestimated disease pathogenesis posteriorly, whereas non-EPI-DWI may be advantageous in this setting.²⁶ Conversely, we managed a patient with an inactive scan who continued to have a $CAS \geq 3$ despite long-term follow-up and systemic immunosuppression, and we suspect this condition was due to unreliable reporting and interpretation of orbital pain.

The other potential limitations of CAS (and other scores based on signs and symptoms) include the absence of a scoring system for highly relevant signs such as diplopia at the initial visit, the inclusion of measurements of ocular motility restriction at follow-up that can be unreliable, and the potential under-reporting of erythema-based signs in darker-skinned individuals.²⁷ The latter point is especially relevant in our patient population in which up to 70% of our patients were not white. Non-EPI-DWI may also have limitations as a disease activity biomarker. It is limited to assessment of ADC to the EOMs only, rather than other tissues implicated in disease pathogenesis such as adipose tissue, and the exact relationship between ADC and disease pathogenesis requires further understanding.

Finally, we found other potential advantages of non-EPI-DWI in our patients that require further validation. By means of ROC analysis, non-EPI-DWI ADC may have the potential to differentiate mild/possible TED unlikely to require immunosuppressive treatment from more severe disease, but this possibility requires validation from larger studies using this imaging technique.

Second, non-EPI-DWI has the potential to gather information on more anatomic variables than CAS alone (up to 8 EOMs), increasing the utility of this technique as a research tool to test hypotheses and explore spatial and temporal patterns of disease. Although not evaluated in this study, it would be of future interest to compare non-EPI-DWI in TED with other causes of EOM enlargement such as lymphoma, metastases, benign tumors, and other inflammatory lesions.

However, despite the potential merits of this technique, it is important to consider the limitations of adjunctive imaging in TED, which include an additional patient visit to the hospital, additional specialist resources, and, as with any investigational test, the potential for false-positive and false-negative findings.

CONCLUSIONS

These results show a positive relationship with CAS and non-EPI-DWI in TED across the spectrum of active disease. To our knowledge, this is the first study demonstrating this relationship using non-EPI-DWI, our preferred technique for imaging soft tissue near the skull base. We also found that our patients fell into 3 activity subgroups, and non-EPI-DWI was particularly useful in aiding clinical decisions with CAS < 3. Our findings also suggest that non-EPI-DWI may have promise as a diagnostic tool, a biomarker of disease activity, and a research tool for hypothesis testing to further our understanding of this challenging and costly disease.

ACKNOWLEDGMENTS

We would like to acknowledge all staff at the Eye and Endocrinology Departments at Central Middlesex Hospital, especially Jennifer Coelho and Drs Koteswara Muralidhara, Camelia Kirillos, Pari Avari, Daniel Darko, Wingmay Kong, and Bernie Colaco. Thanks also to the radiographers and radiology staff at London Northwest University Healthcare National Health Service Trust.

Disclosures: Claire Feeney—UNRELATED: Employment: I have been a Pfizer employee in an unrelated field since June 2018. However, this study was conducted before I started employment and during my PhD with the Imperial College London. Vickie Lee—UNRELATED: Employment: Imperial College Healthcare and London North West Healthcare National Health Service Trusts, Comments: I am employed at these National Health Service institutions as a Consultant Ophthalmic and Oculoplastic Surgeon; Expert Testimony: Eye-Law Chambers, Comments: I undertake medicolegal work in my area of specialist expertise; Other: I work in private practice as a Consultant Ophthalmic and Oculoplastic Surgeon.

REFERENCES

1. Khong JJ, McNab AA, Ebeling PR, et al. Pathogenesis of thyroid eye disease: review and update on molecular mechanisms. *Br J Ophthalmol* 2016;100:142–50 [CrossRef Medline](#)
2. Tanda ML, Piantanida E, Liparulo L, et al. Prevalence and natural history of Graves' orbitopathy in a large series of patients with newly diagnosed Graves' hyperthyroidism seen at a single center. *J Clin Endocrinol Metab* 2013;98:1443–49 [CrossRef Medline](#)
3. Taylor PN, Zhang L, Lee RW, et al. New insights into the pathogenesis and nonsurgical management of Graves' orbitopathy. *Nat Rev Endocrinol* 2020;16:104–16 [CrossRef Medline](#)
4. Rundle FF. Management of exophthalmos and related ocular changes in Graves' disease. *Metabolism* 1957;6:36–48 [Medline](#)
5. Mourits MP, Koornneef L, Wiersinga WM, et al. Clinical criteria for the assessment of disease activity in Graves' ophthalmopathy: a novel approach. *Br J Ophthalmol* 1989;73:639–44 [CrossRef Medline](#)
6. Bartalena L, Baldeschi L, Boboridis K, et al; European Group on Graves' Orbitopathy (EUGOGO). The 2016 European Thyroid Association/European Group on Graves' Orbitopathy Guidelines for the Management of Graves' Orbitopathy. *Eur Thyroid J* 2016;5:9–26 [CrossRef Medline](#)
7. Dolman PJ. Grading severity and activity in thyroid eye disease. *Ophthalmic Plast Reconstr Surg* 2018;34:S34–40 [CrossRef Medline](#)
8. Muller-Forell W, Kahaly GJ. Neuroimaging of Graves' orbitopathy. *Best Pract Res Clin Endocrinol Metab* 2012;26:259–71 [CrossRef Medline](#)
9. Siakallis LC, Uddin JM, Miszkil KA. Imaging investigation of thyroid eye disease. *Ophthalmic Plast Reconstr Surg* 2018;34:S41–51 [CrossRef Medline](#)
10. Cakirer S, Cakirer D, Basak M, et al. Evaluation of extraocular muscles in the edematous phase of Graves' ophthalmopathy on contrast-enhanced fat-suppressed magnetic resonance imaging. *J Comput Assist Tomogr* 2004;28:80–86 [CrossRef Medline](#)
11. Taoka T, Sakamoto M, Nakagawa H, 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 [CrossRef Medline](#)
12. Sillaire I, Ravel A, Dalens H, et al. Graves' ophthalmopathy: usefulness of T2 weighted muscle signal intensity. *J Radiol* 2003;84(2 Pt 1):139–42 [Medline](#)
13. Prummel MF, Gerding MN, Zonneveld FW, et al. The usefulness of quantitative orbital magnetic resonance imaging in Graves' ophthalmopathy. *Clin Endocrinol (Oxf)* 2001;54:205–09 [CrossRef Medline](#)
14. Abdel Razek AA, El-Hadidy M, Moawad ME, et al. Performance of apparent diffusion coefficient of medial and lateral rectus muscles in Graves' orbitopathy. *Neuroradiol J* 2017;30:230–34 [CrossRef Medline](#)
15. McDonald RJ, McDonald JS, Kallmes DF, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology* 2015;275:772–82 [CrossRef Medline](#)
16. Kilicarslan R, Alkan A, Ilhan MM, et al. Graves' ophthalmopathy: the role of diffusion-weighted imaging in detecting involvement of extraocular muscles in early period of disease. *Br J Radiol* 2015;88:20140677 [CrossRef Medline](#)
17. Politi LS, Godi C, Cammarata G, 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 [CrossRef Medline](#)
18. Khemani S, Singh A, Lingam RK, et al. Imaging of postoperative middle ear cholesteatoma. *Clin Radiol* 2011;66:760–67 [CrossRef Medline](#)
19. Lingam RK, Nash R, Majithia A, et al. Non-echoplanar diffusion weighted imaging in the detection of post-operative middle ear cholesteatoma: navigating beyond the pitfalls to find the pearl. *Insights Imaging* 2016;7:669–78 [CrossRef Medline](#)
20. Lingam RK, Bassett P. A meta-analysis on the diagnostic performance of non-echoplanar diffusion-weighted imaging in detecting middle ear cholesteatoma: 10 years on. *Otol Neurotol* 2017;38:521–28 [CrossRef Medline](#)
21. Ritchie AE, Lee V, Feeney C, et al. Using nonechoplanar diffusion-weighted MRI to assess treatment response in active Graves orbitopathy: a novel approach with 2 case reports. *Ophthalmic Plast Reconstr Surg* 2016;32:e67–70 [CrossRef Medline](#)
22. Lingam RK, Mundada P, Lee V. Novel use of non-echo-planar diffusion weighted MRI in monitoring disease activity and treatment response in active Graves' orbitopathy: an initial observational cohort study. *Orbit* 2018;37:325–30 [CrossRef Medline](#)
23. Wong PY, Lingam RK, Pal S, et al. Monitoring progression of 12 cases of non-operated middle ear cholesteatoma with non-echoplanar diffusion weighted magnetic resonance imaging: our experience. *Otol Neurotol* 2016;37:1573–76 [CrossRef Medline](#)
24. Nash R, Lingam RK, Chandrasekharan D, et al. Does non-echo-planar diffusion-weighted magnetic resonance imaging have a role in assisting the clinical diagnosis of cholesteatoma in selected cases? *J Laryngol Otol* 2018;132:207–13 [CrossRef Medline](#)
25. Patel B, Hall A, Lingam R, et al. Using non-echoplanar diffusion weighted MRI in detecting cholesteatoma following canal wall down mastoidectomy: our experience with 20 patient episodes. *J Int Adv Otol* 2018;14:263–66 [CrossRef Medline](#)
26. Smith TJ. Pathogenesis of Graves' orbitopathy: a 2010 update. *J Endocrinol Invest* 2010;33:414–21 [CrossRef Medline](#)
27. Mawn LA, Dolman PJ, Kazim M, et al. Soft tissue metrics in thyroid eye disease: an International Thyroid Eye Disease Society reliability study. *Ophthalmic Plast Reconstr Surg* 2018;34:544–46 [CrossRef Medline](#)