Novel Thermal Imaging Analysis Technique for Detecting Inflammation in Thyroid Eye Disease

Costanzo Di Maria, John Allen, Jane Dickinson, Christopher Neoh, and Petros Perros

Microvascular Diagnostics, Regional Medical Physics Department (C.D., J.A.) and Department of Ophthalmology (J.D., C.N.), Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom; Institute of Genetic Medicine (P.P.), International Centre for Life and Institute of Cellular Medicine (C.D., J.A.), Medical School, Newcastle University, Newcastle upon Tyne, United Kingdom

Context: The disease phase in thyroid eye disease (TED) is commonly assessed by clinical investigation of cardinal signs of inflammation and using the clinical activity score (CAS). Although CAS is the current gold standard, the clinical assessment would benefit if a more objective tool were available.

Objective: The aim of this work was to explore the clinical value of a novel thermal imaging analysis technique to objectively quantify the thermal characteristics of the eye and peri-orbital region and determine the disease phase in TED.

Design: This was a cross-sectional study comparing consecutive patients with active TED (CAS ≥3/7) attending a tertiary center, with a group of consecutive patients with inactive TED (CAS <3).

Patients: Thermal images were acquired from 30 TED patients, 17 with active disease and 13 with inactive disease.

Interventions: Patients underwent standard ophthalmological clinical assessments and thermal imaging.

Main Outcome Measures: Five novel thermal eye parameters (TEP) were developed to quantify the thermal characteristics of the eyes in terms of the highest level of inflammation (TEP1), overall level of inflammation (TEP2), right-left asymmetry in the level of inflammation (TEP3), maximum temperature variability across the eyes (TEP4), and right-left asymmetry in the temperature variability (TEP5).

Results: All five TEP were increased in active TED. TEP1 gave the largest accuracy (77%) at separating the two groups, with 65% sensitivity and 92% specificity. A statistical model combining all five parameters increased the overall accuracy, compared to using only one parameter, to 93% (94% sensitivity and 92% specificity). All five of the parameters were also found to be increased in patients with chemosis compared to those without.

Conclusions: The potential diagnostic value of this novel thermal imaging analysis technique has been demonstrated. Further investigation on a larger group of patients is necessary to confirm these results. (J Clin Endocrinol Metab 99: 4600–4606, 2014)

Thyroid eye disease (TED) is an inflammatory condition which can affect the eyes of patients with autoimmune thyroid disease (1) and leads to significant impairment of the quality of life (QOL) (2, 3). Its incidence is reported to be 42 new cases a year for every million population (4). Useful treatments are comprised of high-dose steroids and other immuno-suppressive drugs, orbital irradiation, and surgery. Choice of treatment is highly dependent on accurate assessment of disease phase, as medical treatments are effective only during the initial
inflammatory phase (“active disease”), while most of the surgical treatments are appropriate only when this phase has subsided (“inactive disease”) (5). The disease phase is commonly assessed clinically, by scoring for the presence or the absence of cardinal inflammatory symptoms and signs to form the clinical activity score (CAS) (6). However, this approach is subjective, relying on the expertise of the clinician, and interpretation of symptoms from the patient, and does not predict response to intervention with high accuracy. In order to develop a more objective assessment tool, several imaging techniques have been investigated, including ultrasound (7, 8), X-ray computer tomography (9, 10), magnetic resonance imaging (MRI) (11, 12), radio-labeled octreotide scanning, gallium scanning (13), 99mTc-DTPA orbital SPECT (14), or a combination of these technologies (15). However, these techniques require an expert operator, or the use of ionizing radiation and most are expensive. Additionally, some require contrast agents, which are not without risk. This pilot study therefore examined medical thermal imaging (TI) (16, 17) for its capability for being a safer and operator-independent technology for the objective assessment of disease phase in TED.

TI is a passive measurement technique which does not require contact with the patient (noninvasive and noncontact) and does not utilize ionizing radiation or contrast agents. A thermal camera collects the infrared radiation spontaneously emitted by the human body and provides information relating to the surface temperature of tissues in the form of a temperature map (thermogram) of the area being imaged. Bespoke algorithms can then be developed to select specific regions of interest, quantify the thermograms, and derive objective measurements. Thermography has already been applied with some success in the diagnosis and assessment of other medical conditions, including Raynaud’s disease and systemic sclerosis (18), restless legs syndrome (19), and the function of renal fistulae (20). Preliminary results on its application in thyroid eye disease have been reported by our group (21). Other studies also used TI to assess the response of patients with thyroid eye disease to methylprednisolone pulse therapy with encouraging results (22, 23).

The aim of this work was to explore the value of a novel thermal imaging analysis technique to objectively quantify the thermal characteristics of the eyes in TED, using CAS as the gold standard.

Materials and Methods

Patients and clinical assessment

This was a cross-sectional study which enrolled 30 consecutive patients with thyroid eye disease attending our tertiary center. Based on recent evidence on the utility of thermography (22, 23), this test was included in the routine clinical assessment of these patients. The patients were clinically assessed by an ophthalmologist using the European Group On Graves’ Orbitopathy (EUGOGO) protocol (24). Demographic and clinical data relevant to TED were collected, including age, gender, smoking habits, and duration of the disease. Serum free tri-iodothyronine (FT3) (normal range 3.5–6.5 pmol/L), free thyroxine (FT4) (normal range 9.5–21.5 pmol/L), and thyroid stimulating hormone (TSH) (normal range 0.3–4.7 mU/L) were measured using electrochemiluminescence technology (Modular E170, Roche Diagnostics, 2-site sandwich assay for TSH, competitive assay for free thyroid hormones). TSH receptor antibody (TBI) levels (normal <1.5 U/L) were measured by a coated tube radioimmunoassay (RIA) kit (RSR Ltd.).

Inflammation was graded using a photographic atlas (5) based on the system derived from the cardinal features of inflammation. This approach consists of assessing the presence or the absence of seven different features: spontaneous retro-bulbar pain; pain with eye movement; eyelid erythema; eyelid swelling; conjunctival injection; chemosis; and caruncle or plical swelling. One point is assigned for each feature present and their sum forms the clinical activity score for each patient (6). Patients with CAS ≥3 out of 7 were considered to have active disease (25), while patients scoring <3 were considered to be inactive.

Thermal imaging measurements

Patients were asked to follow pretest instructions for at least 3 h prior to thermal imaging, avoiding facial cosmetics or talc powder, and avoiding smoking and caffeine. Imaging was performed in a microvascular imaging facility using a cooled room inflammation study protocol with ambient temperature fixed at 18°C. This room was specifically designed for optimal control of a stable uniform temperature, with minimal drafts.

Patients were rested in the waiting area just outside the facility for at least 20 min and then entered the cooled room and acclimatized for 10 min while lying supine on a measurement couch. A FLIR SC300 thermal imaging system (FLIR Systems U.K.) was used to acquire thermal images of the face and eyes (Figure 1) with the patient supine and relaxed, and ceiling lights dimmed for comfort. The skin emissivity was set to 0.98 and the thermal camera was powered up for at least 2 h prior to measurement (26). Multiple thermal images were recorded by an expert operator (JA), over a period no longer than 10 min, at a time when the patient was deemed to be in focus, still, not blinking, facially relaxed, and in the absence of excessive lacrimation. Thermal imaging assessment took place within a median of 27 days (lower quartile 19 and upper quartile 35) from the clinical assessment.

Thermal imaging analysis

Thermal imaging analysis was performed off-line by utilizing dedicated FLIR ThermaCAM Researcher 2001 software (FLIR Systems U.K.).

From the recorded images, the one with the best quality was selected for subsequent analysis by an expert in thermal image analysis (CDiM), different and independently from the one who performed the measurement. For each eye on the thermogram, six anatomical regions of interest (ROI) were identified as illustrated in Figure 2: (1) lateral conjunctiva, (2) cornea, (3) medial conjunctiva, (4) caruncle, (5) upper eyelid, and (6) lower eyelid. In addition, one ROI was selected as the thermal reference for each eye, namely, the lateral orbit (LO). The mean temperature over each ROI was normalized to the reference site to compen-
Sate for differences in basal body temperature between subjects, as shown in Eq. (1). This gave six normalized values (NV) for each eye,

\[ NV_i(\%) = \frac{ROI_i - LO}{LO} \times 100, \quad i = 1, \ldots, 6 \]  

(1)

The overall mean temperature \( T_m(\%) \) for each eye was calculated as the average of the six normalized values [Eq. (2)],

\[ T_m(\%) = \frac{1}{6} \sum_{i=1}^{6} NV_i \]  

(2)

The overall variability of the temperature distribution \( T_{std}(\%) \) across each eye was quantified using the standard deviation (std) of the six normalized values [Eq. (3)],

\[ T_{std}(\%) = \sqrt{\frac{1}{6} \sum_{i=1}^{6} (NV_i - T_m)^2} \]  

(3)

The values obtained from Eqs. (1)–(3) were then further processed to produce five thermal eye parameters (TEP) with the aim of characterizing the thermal properties of the eyes for each subject.

- \( TEP_1 \) denoted the maximum level of inflammation in each subject. This was taken as the highest mean temperature \( T_m \) of either eye.
- \( TEP_2 \) denoted the overall level of inflammation in each patient and was calculated as the average of the mean temperature \( T_m \) of the right and left eye.
- \( TEP_3 \) represented the overall asymmetry in the temperature characteristics between the right and left eyes. It was calculated as the averaged right-left absolute difference in the normalized value for each ROI.
- \( TEP_4 \) quantified the maximum temperature variability across the eyes and was taken as the maximum \( T_{std} \) of either eye.
- \( TEP_5 \) quantified the right-left asymmetry in the temperature variability and was calculated as the absolute right-left difference in \( T_{std} \).

In order to improve the clinical performance of each of the thermal parameters when considered individually, a binary logistic model (BLM) that combined all five TEP was also developed using binary logistic regression (BLR) (27). BLR used the \( n = 5 \) thermal eye parameters to fit the \( n + 1 \) \( b \), coefficients of the logit function \( z \) [Eq. (4)]. Once the coefficients have been fitted, the logistic function \( f(z) \) can be used to calculate the predicted probability (ranging from 0 to 1) for each patient of having the active disease (which is considered as the positive state here). A cut-off value can then properly be set on the predicted probability in order to optimize the classification performance,

\[ f(z) = \frac{1}{1 + e^{-z}}, \quad z = b_0 + \sum_{i=1}^{n} b_i \cdot TEP_i \]  

(4)

**Statistical analysis**

Statistical analysis was carried out with SPSS Statistics (version 20, IBM Corp.) and statistical plots were produced with SigmaPlot (version 11, Systat Software Inc.).

Data were summarized using nonparametric statistics, ie, median value, and lower and upper quartile. The significance of the difference between the two groups was tested by nonparametric tests, the Mann-Whitney U-test when comparing individual values, and the Fisher’s exact test when comparing proportions. Correlations were assessed using Spearman correlation analysis. A \( P \) value <0.05 was considered to represent statistical significance. Discrimination ability was evaluated by receiver operating characteristic (ROC) analysis and quantified by its area under the curve (AUC). The best cut-off value was chosen in order to optimize the Youden’s \( J \)-statistics, ie, the sum of sensitivity and specificity.
specificity (28). Having defined the active disease phase as the positive state, the clinical performance of each parameter with the selected cut-off value was derived from a contingency table and summarized in terms of sensitivity (Se), specificity (Sp), and accuracy. A 95% confidence interval (CI) for A was also calculated as the CI of the proportion (29). If \( r \) individuals out of a sample of size \( n \) experience a specific outcome, then the sample proportion is \( \frac{r}{n} \). The standard error (S.E.) of this proportion is given by Eq. (5), and its 95% CI is from \( p - 1.96 \times \text{S.E.} \) to \( p + 1.96 \times \text{S.E.} \).

\[
\text{S.E.} = \sqrt{\frac{p(1-p)}{n}}
\]

### Results

#### Patients

In total, 5 male and 25 female adults were studied in the age range 21–85 years old (median 54 years old, lower quartile 43 and upper quartile 62). Using a cut-off of ≥3/7, the CAS classified 17 subjects as being in the active disease phase and 13 as inactive. Table 1 summarizes demographics and clinical variables for the active and inactive disease groups. There were no significant differences between the two groups for age or gender, although most of the patients were female with a ratio 1:5 (as was expected from the typical prevalence of this condition). The group with active disease had significantly shorter duration of TED (median 9 months vs 17 months, \( P < .05 \)). In the active group there were 7 subjects with chemosis (41%) whereas there were none (0%) in the inactive group (\( P = .01 \)). There were no significant differences for all other clinical variables.

#### Thermal eye parameters

Figure 3 gives an example of thermal images showing an active TED patient with visibly increased temperature of the eye and peri-orbital region [Figure 3(a)] as compared to a TED patient with inactive disease [Figure 3(b)].

Values obtained for the five TEP are summarized in Table 2. The ac-

---

**Table 1. Summary of Group Demographic and Clinical Data**

<table>
<thead>
<tr>
<th></th>
<th>Active</th>
<th>Inactive</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>51 (39–60)</td>
<td>57 (47–63)</td>
<td>NS</td>
</tr>
<tr>
<td>Gender (males/females)</td>
<td>5/12</td>
<td>0/13</td>
<td>NS</td>
</tr>
<tr>
<td>Smoker (yes/no)</td>
<td>7/10</td>
<td>4/9</td>
<td>NS</td>
</tr>
<tr>
<td>Duration of TED (months)</td>
<td>9.0 (5.0–14.5)</td>
<td>17.0 (9.0–102.0)</td>
<td>(&lt;0.05)</td>
</tr>
<tr>
<td>CAS</td>
<td>4.0 (3.0–4.0)</td>
<td>1.0 (0.0–2.0)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>Range</td>
<td>3–6</td>
<td>0–2</td>
<td></td>
</tr>
<tr>
<td>Proptosis (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right eye</td>
<td>22.0 (19.8–23.3)</td>
<td>21.0 (20.5–21.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Left eye</td>
<td>21.5 (20.0–23.0)</td>
<td>21.0 (19.0–22.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Palpebral aperture (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right eye</td>
<td>10.5 (9.8–12.5)</td>
<td>10.0 (10.0–13.0)</td>
<td>NS</td>
</tr>
<tr>
<td>Left eye</td>
<td>11.0 (10.3–12.0)</td>
<td>10.0 (9.0–12.0)</td>
<td>NS</td>
</tr>
<tr>
<td>Diplopia (yes/no)</td>
<td>7/10</td>
<td>0/13</td>
<td></td>
</tr>
<tr>
<td>Chemosis (yes/no)</td>
<td></td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>fT3 (pmol/L)</td>
<td>4.6 (3.8–5.5)</td>
<td>4.4 (4.2–5.0)</td>
<td>NS</td>
</tr>
<tr>
<td>fT4 (pmol/L)</td>
<td>16.7 (14.7–23.8)</td>
<td>20.6 (16.3–25.1)</td>
<td>NS</td>
</tr>
<tr>
<td>TSH (mU/L)</td>
<td>0.08 (0.05–1.81)</td>
<td>0.64 (0.06–1.42)</td>
<td>NS</td>
</tr>
<tr>
<td>TBII (U/L)</td>
<td>3.8 (1.6–23.2)</td>
<td>1.6 (1.2–14.4)</td>
<td>NS</td>
</tr>
</tbody>
</table>

Abbreviation: NS, not significant. Values are reported as median (lower quartile – upper quartile) or counts. The \( P \) value indicates if any statistically significant difference was found comparing active patients against inactive patients. Statistically significant \( P \) values are in bold.

---

Figure 3. Example thermal images from (a) a patient with active TED and (b) a patient with inactive TED. The thermal image for the patient with active TED (a) shows visibly increased temperature of the eye and peri-orbital region compared to the patient with inactive TED (b). Notice that the two images are shown in the same temperature scale to facilitate visual comparison.
The results of Se, Sp, and A were calculated using the specified cut-off value for each parameter. The best performance is in bold.

Table 2. Statistical Summary of the Five Thermal Eye Parameters

<table>
<thead>
<tr>
<th></th>
<th>Active</th>
<th>Inactive</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEP1</td>
<td>5.56 (2.76–6.77)</td>
<td>2.81 (2.23–3.78)</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>TEP2</td>
<td>4.43 (2.39–5.14)</td>
<td>2.49 (1.46–3.31)</td>
<td>0.01</td>
</tr>
<tr>
<td>TEP3</td>
<td>1.76 (1.15–2.85)</td>
<td>1.22 (0.93–1.76)</td>
<td>NS</td>
</tr>
<tr>
<td>TEP4</td>
<td>1.97 (1.62–2.50)</td>
<td>1.79 (1.36–2.35)</td>
<td>NS</td>
</tr>
<tr>
<td>TEP5</td>
<td>0.51 (0.28–0.86)</td>
<td>0.26 (0.07–0.69)</td>
<td>NS</td>
</tr>
</tbody>
</table>

Abbreviation: NS, not significant. Values are reported as median (lower quartile – upper quartile). The P value indicates if any statistically significant difference was found comparing active against inactive patients. Statistically significant P values are in bold.

tive group had both significantly higher TEP1 (median 5.56% vs inactive 2.81%, P < .05) and TEP2 (4.43% vs inactive 2.49%, P = .01). ROC analysis for these two parameters calculated an AUC of 0.73 and 0.76, respectively. The binary logistic model combining all five TEP improved these results giving an AUC of 0.96. Overall, the most accurate parameter for identifying the disease phase was TEP1, with an accuracy of 77% using a cut-off value of 4.80%. Classification performance for all other parameters derived from the thermograms is also given in Table 3. The BLM improved this result giving an overall accuracy of 93% using a cut-off value of 0.57 for the BLR predicted probability (Table 3).

All five thermal eye parameters were increased in patients with chemosis as compared to the patients without (figure 4) and the parameter TEP4 approached statistical significance with a P value of 0.06 in this case. Although correlations with other clinical variables listed in Table 1 were not statistically significant, TEP1 and TEP2 were weakly but significantly correlated with CAS values with a Spearman ρ correlation coefficient of 0.386 (P < .05) and 0.469 (P < .01), respectively. The other three TEP did not correlate significantly with CAS. The predicted probabilities obtained from the BLM also correlated significantly with CAS (ρ = 0.632, P < .001).

Discussion

TED patients are usually treated by medication or surgical intervention depending on clinical evaluation of disease phase. Therefore, accurate assessment of disease phase is of paramount importance to inform the most appropriate treatment. As TED is an inflammatory condition, then systems to score clinical signs of inflammation were developed, and the clinical activity score has been widely adopted for this assessment. However, its shortcomings have become increasingly apparent: it does not correlate well with disease duration as would be expected in a condition with one acute episode; not all patients with progressive, short duration TED show inflammatory signs (negative CAS but progression suggesting active disease); and conversely, patients with high CAS sometimes show persistent inflammatory signs for years that may reflect venous congestion alone. Objective measurements, such as ultrasound, MRI, and serum markers have so far shown not to be able to identify active TED more accurately than CAS, although a complex prediction model amalgamating many parameters was more accurate in predicting both response and no-response to radiotherapy (15). Hence, there is a compelling need to explore other modes of assessment that could be more accurate in predicting a response to medical intervention, the practical goal of identifying activity, or alternatively excluding activity and therefore determining the earliest time for safe rehabilitation surgery. In seeking to find new assessment tools, it remains valid to compare such tools with CAS as it remains in worldwide use in both clinical management and research. Indeed, studies on medical intervention routinely use CAS for patient selection (30), as a primary outcome measure (31), and additionally it has also been shown to correlate with orbital TSH receptor expression in patients classified as active or inactive by other parameters (32). Previous studies using thermal imaging also utilized CAS as the gold standard to quantify changes to treatment (22, 23). The latter two studies only quantified the temperatures of the individual regions of interest. This can provide ambiguous results with a patient having normal temperature for one ROI and elevated temperature for another ROI leading to difficult clinical interpretation of disease phase. In this study, we further processed these raw temperature measurements in order to define general clin-
ical characteristics of inflammation, which could be easier and more objective to interpret.

We employed medical thermal imaging assessment in our standard cooled room microvascular protocol to investigate its diagnostic value in identifying TED phase. Consistency of the measurement conditions across patients was achieved by the use of a temperature-controlled thermal imaging facility. Thermal images were acquired by a single expert operator in the absence of possible confounders such as patient movement, blinking, or lacrimation. We built on the current state of the art and successfully explored novel thermal imaging processing methods to provide an objective quantification of the thermograms. The thermal characteristics of the eyes are relatively complex, but we developed five thermal eye parameters to summarize the ocular characteristics relevant to identifying inflammation. All patients participating in this study had normal serum fT3 and fT4 levels and, therefore, can be considered to be biochemically euthyroid. All five TEP were increased in the group with active disease (CAS ≥3), two significantly so: TEP1 (quantifying the highest degree of inflammation of either eye) and TEP2 (which averaged the overall inflammation between the two eyes). TEP1 gave the best classification accuracy of 77%; TEP2 gave the highest ROC AUC (0.76). The BLM which included all five thermal eye parameters improved both ROC AUC (0.96) and overall classification accuracy (93%). This final result is encouraging, suggesting that combining more parameters could possibly further improve classification accuracy. It is important to stress that this multivariate model only used parameters obtained objectively from one single thermal image by utilizing our novel imaging analysis technique, and not by amalgamating parameters from diverse tests and clinical assessments. All five TEP were increased in the group of patients with chemosis with a trend towards statistical significance for TEP4 (P = .06). Furthermore, TEP1 and TEP2 were weakly but significantly correlated with CAS values with ρ = 0.386 (P < .05) and ρ = 0.469 (P < .01), respectively, suggesting a possible link between this thermal imaging assessment and clinical disease activity.

Thermography is a sensitive, noninvasive technique which can be less expensive than other traditional medical imaging techniques (eg, MRI or computer tomography), once the equipment and operator expertise is established. It is important though to have a facility with controlled temperature environment to guarantee consistency of the measurement conditions. The role of other possible confounders, such as coexistent unrelated ocular inflammation or even inadvertent eye rubbing, also needs to be taken into account and possibly controlled as these may affect the thermal image and potentially yield to misleading results. It is important, therefore, that a strict protocol is followed and measurement performed by an experienced operator.

Future work is now required to validate these results prospectively in a larger group of patients and to assess the value of using thermal imaging to predict response to therapy independently from the clinical activity score.

Conclusion

This work has confirmed previous reports of the potential utility of medical thermal imaging in the assessment of patients with thyroid eye disease. A novel thermal imaging processing algorithm has been successfully developed in this pilot and its possible diagnostic value has been demonstrated. Further investigation on a larger group of patients is necessary to confirm these results and to test the possible value of TEP as predictors of response to medical treatments.

Acknowledgments

Address all correspondence and requests for reprints to: Costanzo Di Maria, Microvascular Diagnostics, Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne, NE7 7DN, United Kingdom. E-mail: costanzo.dimaria@nuth.nhs.uk.

A contribution to the thermal imaging measurement system was provided from the “Endocrine Research Fund.”

Disclosure Summary: The authors have nothing to disclose.

References


