Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy
Mourits, M.P.; Prummel, M.F.; Wiersinga, W.M.; Koornneef, L.

Published in:
Clinical endocrinology

DOI:
10.1046/j.1365-2265.1997.2331047.x

Citation for published version (APA):
Clinical activity score as a guide in the management of patients with Graves’ ophthalmopathy

Maarten Ph. Mourits, Mark F. Prummel*, Wilmar M. Wiersinga* and L. Koornneef
Orbital Centre and *Department of Endocrinology, Academic Medical Center, University of Amsterdam, The Netherlands

(Received 10 September 1996; returned for revision 27 October 1996; finally revised 13 November 1996; accepted 25 April 1997)

Summary

OBJECTIVE Approximately 35% of patients with Graves’ ophthalmopathy do not respond to immunosuppressive treatment. A possible explanation for this finding is that only patients with active ophthalmopathy respond to immunosuppressive treatment, whereas patients with fibrotic end stage disease do not. To distinguish between these two groups and to predict the outcome of immunosuppressive treatment, we developed a clinical activity score (CAS) based on four of the five classical signs of inflammation and tested its efficacy in a double-blind, prospective study.

DESIGN, PATIENTS AND MEASUREMENTS The CAS was determined by an ophthalmologist before, on the day of, and after the start of either oral prednisone or retrobulbar irradiation in 43 patients with moderate to severe Graves’ ophthalmopathy. The therapeutic outcome was determined by a second ophthalmologist unaware of the CAS stores given. Success of treatment was defined as an improvement in NOSPECS class or grade.

RESULTS Responders (22) and non-responders (21) did not differ in age, sex, duration or severity of their Graves’ ophthalmopathy. The pretreatment CAS, however, was significantly higher in responders than in non-responders. Twelve of 22 responders and three of 21 non-responders had a CAS ≥ 4 (80% vs 36%; P < 0.01). Using this CAS cut-off point, the accuracy of CAS in predicting the therapeutic outcome was: specificity 86%, sensitivity 55%, positive predictive value 80%, negative predictive value 64%. Patients with a CAS ≥ 4 had a similar duration of Graves’ ophthalmopathy as patients with a CAS < 4.

CONCLUSIONS The clinical activity score has a high predictive value for the outcome of immunosuppressive treatment in Graves’ ophthalmopathy. Disease activity, and not disease duration, is the prime determinant of therapeutic outcome.

Graves’ ophthalmopathy is a disease characterized by enlargement of the extraocular muscles and an increase in retrobulbar fat (Bahn & Heufelder, 1993). As a consequence, the eyelids become swollen and the eyes proptotic. Orbital fibroblasts are thought to play a crucial role in the pathogenesis by their ability to produce hydrolphilic glycosaminoglycans. The presence of activated CD4+ lymphocytes in the orbits of patients with Graves’ ophthalmopathy and the interactions of these cells with retrobulbar fibroblasts together with other supporting findings has led to the assumption that Graves’ ophthalmopathy is an organ-specific autoimmune disease and can therefore be treated by means that suppress the immune process (Burch & Wartofsky, 1993). One such treatment is administration of corticosteroids either intravenously or orally. Another is retrobulbar external irradiation in a total dosage of 20 Gy, which inhibits but does not kill the fibroblasts present (Kahaly & Beyer, 1988). Several studies have shown that both treatment modalities proved effective in a majority of patients (Wiersinga et al., 1988; Prummel et al., 1989, 1993a). However, a substantial number did not respond to these treatments (Mourits et al., 1989). As Graves’ ophthalmopathy appears to be biphasic, with an initial phase of active and progressive disease followed by a ‘burnt out’ phase in which alterations such as proptosis and motility restrictions remain stable, we hypothesized that immunosuppression would be beneficial in only the early, active phase of the disease. This assumption is generally accepted, but never proved. We, therefore, tried to develop an ‘activity score’ in patients with Graves’ ophthalmopathy, that could predict which patients would benefit from immunosuppressive treatment and which would not. This clinical activity score (CAS) is based on four of the five well-known classical signs of inflammation (pain, redness, swelling and impaired function) and consists of 10 items (Table 1). For each item present, one point is given. Each item has the same weight. The sum of these points is the CAS (range 0–10). In a retrospective pilot study we found a positive
relationship between this CAS and the outcome of immuno-
suppressive treatment (Mourits et al., 1989). The present
study was undertaken to test the usefulness of the CAS in a
prospective, double-blind trial in which patients with moderate
to severe Graves’ ophthalmopathy were treated with either
prednisone or with retrobulbar irradiation.

Methods

Patients

We studied patients with moderately severe Graves’ ophthal-
mopathy, who were between 20 and 75 years old. All had given
their informed consent for the study. All had been euthyroid for
at least 3 months, as indicated by the presence of normal plasma
concentrations of T4 and T3, and by TSH concentrations that
were within or below the normal range (normal ranges: T4,
70–150 nmol/l; T3, 1:30–2:45 nmol/l; TSH, 0:4–4:0 mU/l). Most patients were on antithyroid drugs in combination with T4
supplementation (n = 28). Five patients used T4 after previous
ablative treatment, six were in remission of previous Graves’
hyperthyroidism, while in four patients no evidence for Graves’
thyroid disease was found. The diagnosis of Graves’ ophthal-
mopathy was based on the presence of the typical clinical
features in combination with enlarged extraocular eye muscles
on a coronal CT scan. We included only patients with
moderately severe ophthalmopathy, defined as having one or
more of the following NOSPECS categories: class 2, grade c;
class 3, grades abc; class 4, grades abc and class 6, grade a

(Table 2; Werner, 1977). Patients with diabetes mellitus were
not included, neither were patients with contraindications for
corticosteroids or external beam irradiation. Patients with
rapidly decreasing visual functions or other signs of optic
nerve compression by apical crowding were not included. Prior
treatment other than eyedrops, gels or ointments excluded
patients from participation. Patients with insufficient follow-up,
or who developed abnormal thyroid function during the study
were excluded from the analysis and replaced. The study was
approved by the Medical Ethics Committee of the University of
Amsterdam.

Treatment

The patients participated in a clinical trial comparing oral
prednisone and radiotherapy (Prummel et al., 1993a), and were
randomly assigned to receive oral prednisone or retrobulbar
external beam irradiation. Prednisone was given in decreasing

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The 10 items of the clinical activity score (CAS)</th>
</tr>
</thead>
</table>
| Pain    | 1 Painful, oppressive feeling on or behind the
globe, during the last 4 weeks |
|         | 2 Pain on attempted up, side or down gaze,
during the last 4 weeks |
| Redness | 3 Redness of the eyelid(s) |
|         | 4 Diffuse redness of the conjunctiva, covering
   at least one quadrant |
| Swelling| 5 Swelling of the eyelid(s) |
|         | 6 Chemosis |
|         | 7 Swollen caruncle |
|         | 8 Increase of proptosis of ≥2 mm during a
   period of 1–3 months |
| Impaired function | 9 Decrease of eye movements in any direction
   ≥5° during a period of 1–3 months |
|         | 10 Decrease of visual acuity of ≥1 line(s) on
   the Snellen chart (using a pinhole) during
   a period of 1–3 months |

For each item present, 1 point is given. The sum of these points is the
CAS, e.g. a CAS of 6 means that six items were present, regardless of
which items.

| Table 2 | The NOSPECS classification of ocular changes in Graves’
ophthalmopathy* |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>Grade</td>
</tr>
<tr>
<td>0</td>
<td>No signs or symptoms</td>
</tr>
<tr>
<td>1</td>
<td>Only signs</td>
</tr>
</tbody>
</table>
| 2       | Soft tissue involvement, with symptoms and signs
   | o absent |
|         | a minimal |
|         | b moderate |
|         | c marked |
| 3       | Proptosis
   | o <23 mm |
|         | a 23–24 mm |
|         | b 25–27 mm |
|         | c ≥28 mm |
| 4       | Extraocular muscle involvement
   | o absent |
|         | a limitation of motion in extremes of gaze |
|         | b evident restriction of motion |
|         | c fixation of a globe or globes |
| 5       | Corneal involvement
   | o absent |
|         | a stippling of cornea |
|         | b ulceration |
|         | c clouding, necrosis, perforation |
| 6       | Sight loss (visual acuity†)
   | o >0:67 |
|         | a 0:67–0:33 |
|         | b 0:33–0:10 |
|         | c <0:10 |

* Modified slightly from the classification of Werner (1977).
† Expressed as a decimal, e.g. 20/20 = 1:0, 20/30 = 0:67, 20/60 =
0:33, 20/200 = 0:10.
Clinical activity score in Graves’ ophthalmopathy

To assess the disease activity, the CAS was used (Table 1). To assess the disease severity and the outcome of treatment, the NOSPECS classification was used (Table 2).

The CAS was determined in all patients 2 days before (T₀) of the start of treatment and 24 weeks (T₂₄) later by an ophthalmologist (L.K.) who was otherwise not involved in the diagnosis or treatment of these patients. The CAS on T₀ was used to establish the relationship with the outcome of treatment. The CAS on T⁻² and on T₀ were used to calculate the Kappa value, thus determining the intraobserver agreement. The changes between the CAS on T₀ and T₂₄ served to assess the effect of immunosuppressive treatment on the disease activity.

Patients were classified according to the NOSPECS classification by a second ophthalmologist (M.P.M.) on the day of the start of treatment and 4, 12 and 24 weeks later. This second ophthalmologist has no knowledge of the CAS that had been given to the patients. Proptosis was measured with the same Hertel exophthalmometer, and pinhole visual acuity was tested with the Snellen chart and expressed as a decimal (e.g. 20/20 = 1.0; 20/30 = 0.67; 20/60 = 0.33). The ocular motility was described in terms of the NOSPECS classification (e.g. normal, slightly, moderately and severely impaired). In addition, the eye movements in four directions (up, down and sideways) were measured in degrees by an orthoptic assistant using a Maddox cross (Crone, 1973). As the end-point of the study we determined the therapeutic outcome at 24 weeks after start of treatment. A response was defined as a decrease in NOSPECS class or, if this did not occur, in grade from the baseline value. A lack of response was indicated by an absence of change in class or grade, or by treatment failure if an increase in class or grade occurred. If both increases in grades were observed, the change in the highest NOSPECS class determined the therapeutic outcome (Krummel et al., 1989). Because corneal involvement in Graves’ ophthalmopathy usually disappears with local measures and therefore does not necessarily indicate more severe eye disease, NOSPECS class 5 was not used to determine treatment response.

Results

The study group consisted of 43 patients who fulfilled the inclusion criteria. Twenty-two patients were treated with prednisone and 21 received radiotherapy. One patient developed a recurrence of hyperthyroidism during the study period and was replaced. Two patients showed a serious deterioration of their ophthalmopathy at the end of the study period and did not complete the follow-up of 24 weeks. The CAS at the time of their premature withdrawal, however, was determined and their data were included in the analysis.

The therapeutic outcome after 24 weeks, as determined by the change in the highest NOSPECS class, did not differ in the prednisone group (10 of 22 responders) and the radiotherapy group (12 of 21 responders). The total success rate was 51%. Responders and non-responders did not differ in age (43.5 ± 11.9 vs 46.2 ± 11.4 years (mean ± SD), not significant) or sex (20 females/2 males vs 17 females/4 males, not significant). The duration of their ophthalmopathy was similar (median 14 (range 6–120) vs median 12 (range 5–96) months, not significant). Also, the severity as expressed as the Total Eye Score (TES = sum of each NOSPECS class present multiplied by the grade within that class, in which we substituted a, b and c, respectively for 1, 2 and 3) was not different (10.8 ± 4.7 vs 9.8 ± 3.7 (mean ± SD), not significant). However, the disease activity as indicated by the CAS on T₀ was significantly higher in responders than in non-responders (3.4 ± 2.0 vs 2.4 ± 1.2 (mean ± SD), P = 0.05).

The relationship between the CAS on T₀ and the therapeutic outcome is depicted in Fig. 1. Twelve of 22 responders and three of 21 non-responders had a CAS ≥ 4 (80% vs 36%; P < 0.01). Taking this CAS cut-off point, the accuracy of CAS in predicting the therapeutic outcome is given in Table 3. The accuracies of the individual items of the CAS are given
in Table 4. If a CAS cut-off point of 3 was used, this resulted in less acceptable values (positive predictive value 61%, negative predictive value 60%, sensitivity 64% and specificity 57%). Comparing patients with a CAS > 4 and patients with a CAS < 4, we found a similar duration of Graves’ ophthalmopathy (median 17.0 (range 7–120) vs median 12.0 (range 5–96) months, not significant (Table 5)), but a different severity of the ophthalmopathy in both groups (10.9 ± 6.4 vs 8.1 ± 2.9 (mean ± SD), P = 0.001). Twenty-four weeks after the start of treatment, CAS had decreased only in responders (3.4 ± 2.0 to 1.9 ± 1.9 (mean ± SD), P < 0.001), whereas it did not change in non-responders (2.4 ± 1.2 to 2.7 ± 1.9 (mean ± SD), not significant).

The reproducibility of the pretreatment individual CAS items as well as the total CAS (as determined on T⁻² and on T⁰ by calculating the Kappa values) is shown in Table 4. Kappa values varied from 36 (chemosis) to 100 (caruncle swelling, proptosis increase, decrease in motility, decrease in visual acuity). Eyelid swelling was the most prevalent item, whereas caruncle swelling and decreased visual acuity were only observed once. Eyelid swelling appeared to have the lowest predictive value; caruncle swelling, increased proptosis, decreased motility and decreased visual acuity had the highest predictive values.

### Table 3 Accuracy of clinical activity score (CAS) in predicting therapeutic outcome of immunosuppression in Graves’ ophthalmopathy

<table>
<thead>
<tr>
<th>CAS item</th>
<th>Success</th>
<th>No success</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAS ≥4</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>CAS &lt;4</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>21</td>
</tr>
</tbody>
</table>

Positive predictive value, 12 of 15 = 80%; negative predictive value, 18 of 28 = 64%; sensitivity, 12 of 22 = 55%; specificity, 18 of 21 = 86%.

### Table 4 Observed agreement and predictive value of individual items of the clinical activity score (CAS)

<table>
<thead>
<tr>
<th>Incidence</th>
<th>CAS item</th>
<th>T⁰</th>
<th>T⁻²</th>
<th>OA</th>
<th>Ka</th>
<th>PPV</th>
<th>NPV</th>
<th>Sens</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain at rest</td>
<td>53</td>
<td>60</td>
<td>84</td>
<td>67</td>
<td>57</td>
<td>59</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain moving</td>
<td>49</td>
<td>49</td>
<td>86</td>
<td>72</td>
<td>50</td>
<td>52</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red eyelid</td>
<td>23</td>
<td>26</td>
<td>84</td>
<td>57</td>
<td>60</td>
<td>52</td>
<td>27</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Red conjunctiva</td>
<td>28</td>
<td>19</td>
<td>77</td>
<td>36</td>
<td>67</td>
<td>55</td>
<td>36</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Swelling lid</td>
<td>91</td>
<td>74</td>
<td>84</td>
<td>47</td>
<td>51</td>
<td>50</td>
<td>91</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Chemosis</td>
<td>21</td>
<td>26</td>
<td>84</td>
<td>57</td>
<td>78</td>
<td>56</td>
<td>47</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Caruncle</td>
<td>2</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>5</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proptosis</td>
<td>19</td>
<td>19</td>
<td>100</td>
<td>100</td>
<td>58</td>
<td>32</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motility</td>
<td>7</td>
<td>7</td>
<td>100</td>
<td>100</td>
<td>53</td>
<td>14</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual acuity</td>
<td>2</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>5</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total CAS</td>
<td>—</td>
<td>—</td>
<td>88</td>
<td>73</td>
<td>80</td>
<td>64</td>
<td>55</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>

Incidence. In per cent (on day of start of treatment, T⁰, and 2 days before, T⁻²); OA, observed agreement T⁰/T⁻²; Ka, Kappa value; PPV, positive predictive value; NPV, negative predictive value, Sens, sensitivity and Spec, specificity of the 10 items of the CAS and of the total CAS (cut-off point 4) in 43 patients treated for Graves’ ophthalmopathy. For the calculation of the PPVs, NPVs, sensitivity and specificity the CAS items on T⁰ were used. PPVs, etc. were calculated for treatment outcomes expressed in NOSPECS score alteration.

### Discussion

Rundle (1964) has shown that the typical course of Graves’ ophthalmopathy is characterized by an initial phase of rapid progression, reaching its peak after 6–24 months, followed by a plateau phase of varying duration, and then a gradual, often incomplete regression of symptoms. It is conceivable that the signs and symptoms of the initial phase reflect the inflammatory changes induced by the autoimmune process, whereas the white proptotic eyes with upper and lower lid retraction and unchanging motility impairment are expressions of the fibrotic end stage of the disease (Burch & Wartofsky, 1993). As the duration of the various stages of the disease may vary...
canthus over the extraocular muscle insertions. Redness of the orbital tissues, nor the dilated vessels in the lateral ulceration is not what we consider a sign of active inflammation that redness of the conjunctiva as a result of corneal stippling or that this sign is difficult to recognize. It should be emphasized that redness of the conjunctiva had the lowest Kappa value, suggesting this sign’s chance in this study was 51%. The intraobserver variance priori of inflammation. The item ‘heat’ was not incorporated, as impaired function. The CAS is based on these principal signs has been described in terms of pain, redness, swelling, heat and impossible without special instruments.

This prospective, double-blind study confirmed the provisional conclusions of our retrospective pilot study: CAS has a high specificity and high positive predictive value in predicting the therapeutic outcome of immunosuppressive treatment and radiotherapy. Patients with a CAS ≥ 4 had a 80% chance of a favourable treatment effect, whereas the a priori chance in this study was 51%. The intraobserver variance (expressed in Kappa values) appeared very reasonable.

Subjective complaints, such as pain at rest and pain during eye movements, were reasonably consistently revealed. Redness of the conjunctiva had the lowest Kappa value, suggesting that this sign is difficult to recognize. It should be emphasized that redness of the conjunctiva as a result of corneal stippling or ulceration is not what we consider a sign of active inflammation of the orbital tissues, nor the dilated vessels in the lateral canthus over the extraocular muscle insertions. Redness of the conjunctiva as a sign of acute inflammation is a diffuse redness, covering at least one quadrant. The low Kappa values of the sign of redness may further be due to the biological variation. Swelling of the eyelids was diagnosed in almost all patients and was therefore non-specific. It was difficult to discriminate between eyelid swelling caused by oedema and by fat protrusion. However, only the former is thought to be related to the acute inflammatory process. Chemosis, caruncle swelling, increase in proptosis, decrease in motility and visual acuity, although present only in a small minority of patients, were all highly specific for disease activity and therapeutic outcome.

We tried to correlate the CAS without the three last items with the therapeutic outcome, but found no relationship. Therefore, two visits with at least a month in between, are required to assess the CAS. In practice, we do not consider this a major disadvantage as this time can be used to collect data on the patient’s status (thyroid function, CT scan, etc.). More important is the fact that, although very specific, the CAS is not very sensitive. Thirty-six per cent of patients with a CAS < 4 nevertheless showed a favourable outcome on treatment. We have tried to modify the CAS in such a way that some items had a double or triple weight, but this did not result in a more sensitive CAS. Therefore, we believe that the CAS should be used in combination with other parameters of disease activity.

Several laboratory parameters of disease activity in Graves’ ophthalmopathy have been proposed. Kahaly has extensively studied the concentrations of glycosaminoglycans (GAGs) in the urine and plasma of patients with active and inactive, treated and not (yet) treated Graves’ ophthalmopathy (Kahaly et al., 1990, 1994; Schuler et al., 1993). He found a significant

© 1997 Blackwell Science Ltd, Clinical Endocrinology, 47, 9–14
increase in GAGs in both the urine and plasma of patients with Graves’ ophthalmopathy as compared with control groups. This increase was most pronounced in patients described as having active disease (based on our CAS) and subsided after immunosuppressive treatment and radiotherapy. The concentrations of GAGs might therefore be used as a marker of disease activity. These are, however, influenced by several pathological conditions and their significance as a predictor of therapeutic outcome has still to be investigated. Hoh et al. (1994) used the short tau inversion recovery sequence in magnetic resonance imaging to estimate the water content (oedema) of orbital tissues, thus assessing disease activity in Graves’ ophthalmopathy. They found a correlation between high signal intensity ratios and active disease as described in terms of our CAS. The predictability of therapeutic response using this method is currently being investigated. Prummel et al. (1993b) used A-mode ultrasonography to assess disease activity in Graves’ ophthalmopathy and found a positive predictive value of 73% in predicting the outcome of prednisone treatment or radiotherapy, albeit in a very limited number of patients. Also, serum soluble interleukin-2 receptor and serum soluble CD8 concentrations have been proposed, but not tested, as markers of disease activity (Weryha et al., 1991; Balazs et al., 1994).

The advantage of the clinical activity score is that it is easy, cheap and entirely clinical. In patients in whom spontaneous recovery cannot be awaited, treatment usually consists of corticosteroid treatment, orbital irradiation or surgery. Total duration of the ophthalmopathy is not a good guide in deciding which therapy is most appropriate. Early surgery may result in recurrence of symptoms; immunosuppression by steroids or irradiation given in the fibrotic end stage may unnecessarily prolong the disease period. A high clinical activity score helps to select patients who will benefit from immunosuppressive treatment. A low clinical activity score, however, does not exclude favourable therapeutic results. Further studies to improve the sensitivity of the clinical activity score or the predictability or a combination of disease activity markers are therefore required.

References

© 1997 Blackwell Science Ltd, Clinical Endocrinology, 47, 9–14