EULAR response criteria for polymyalgia rheumatica: results of an initiative of the European Collaborating Polyarthritis Rheumatica Group (subcommittee of ESCISIT)

B F Leeb, H A Bird, G Nesher, I Andel, W Hueber, D Logar, C M Montecucco, J Rovensky, J Sautner, M Sonnenblick

Objective: To develop response criteria for polymyalgia rheumatica (PMR) for monitoring treatment and comparing alternative treatments regimens.

Methods: 76 patients, mean (SD) age 68.7 (7.7) years, were enrolled. Corticosteroids, and non-steroidal anti-inflammatory drugs (NSAIDs) were the only drugs allowed during the observation period. Erythrocyte sedimentation rate (ESR), C reactive protein (CRP), α2 globulin, serum iron, pain, physician’s global assessment (PGA), morning stiffness (MST), muscle tenderness (MT), myalgia, and the elevation of upper limbs (EUL) were determined regularly. The daily corticosteroid and NSAID doses as the corticosteroid response time were recorded. To ensure evaluation of an adequate number of patients (n = 57) week 24 was chosen for final analysis.

Results: ESR, CRP, α2 globulin, pain, PGA, MST, myalgia, MT, and EUL showed significant improvement (p < 0.0001) at week 24 compared with week 0. Multiple regression analysis showed that changes of ESR (p = 0.08), CRP (p = 0.41), α2 globulin (p = 0.13), MST (p = 0.1), and MT (p = 0.07) were independent of pain, but myalgia (p < 0.001) and EUL (p = 0.003) were pain dependent. Consequently, a core set of PMR response criteria, comprising ESR or CRP, pain, PGA, MST, and EUL was established. Assessment of treatment responses with this core set resulted in 90%, 70%, 50%, and 20% improvement in 31/57 (54%), 46/57 (81%), 51/57 (89%), and 54/57 (95%) of the patients, respectively.

Conclusion: These PMR response criteria are a promising tool for better monitoring of disease activity and treatment in PMR. It is proposed that these criteria should be used in clinical trials in the near future to explore alternative treatment options for PMR.

Polymyalgia rheumatica (PMR) constitutes a common rheumatic disorder in the elderly population, affecting from 0.1 to 0.5% of people older than 50 years. However, it may also occur in younger people, but the incidence is lower. The diagnosis is based upon a clinical syndrome, consisting of pain and stiffness in the shoulder and pelvic girdle, muscle tenderness of the upper and lower limbs, non-specific somatic complaints, frequently occurring fever, weight loss, and fatigue. Moreover the acute phase response measured by erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) is greatly increased. However, PMR may also exist with a low ESR and CRP. Whether a low acute phase response indicates less severity and better prognosis is not definitely known.

Generally, corticosteroid treatment leads to a rapid and dramatic improvement of patients’ complaints and returns them to previous functional status. Despite severe complications caused by giant cell arteritis, which is often associated with PMR, the overall prognosis of the disease can be regarded as good, although prolonged corticosteroid treatment, sometimes for several years, may be necessary to maintain clinical improvement. Long term corticosteroid use can be associated with various adverse events, with the induction of osteoporosis, diabetes, or infection among the worst. Therefore other drugs such as methotrexate or azathioprine have been suggested as corticosteroid sparing agents, and have been investigated in a few clinical trials with contradictory outcomes in many cases owing to the lack of comparable response criteria.

In 1997 a European collaborative PMR group was initiated as a subcommittee of ESCISIT (EULAR Standing Committee on Clinical Trials Including Therapeutic Trials) with the objective to validate existing diagnostic criteria for PMR and establish PMR response criteria which could be used in future clinical trials. Here we report the development, definition, and assessment of PMR response criteria in a European/Israeli cohort of patients, and the validation of these criteria in an additional group of patients with this disease.

PATIENTS AND METHODS

The European collaborative PMR group consists of eight centres representing all parts of the continent (Leeds, UK; Vienna, Austria; Ljubljana, Slovenia; Kaunas, Lithuania; Pavia, Italy; Tartu, Estonia; Piestany, Slovak Republic, and Israel, Jerusalem). Two hundred and thirteen patients were enrolled into the diagnostic criteria validation study. All patients from this study with proven and newly diagnosed PMR who were available for follow up were subsequently included in the PMR response criteria study (henceforth called the European patient group).

After informed consent was obtained at study entry, a full medical history was obtained and a detailed physical examination was performed, including assessment of all signs and symptoms that were previously suggested for the

Abbreviations: CRP, C reactive protein; ESCISIT, EULAR Standing Committee on Clinical Trials Including Therapeutic Trials; ESR, erythrocyte sedimentation rate; MST, morning stiffness; MT, muscle tenderness; NSAIDs, non-steroidal anti-inflammatory drugs; PGA, physician’s global assessment; PMR, polymyalgia rheumatica; VAS, visual analogue scale
diagnosis of PMR. The Bird/Wood 1979 criteria best identified patients from this group of 213 considered to have PMR by 10 experienced investigators from across Europe. The identification range was found to be between 99.5% and 67.8% for the four sets of criteria applied (paper in preparation). The parameters listed in table 1 were also determined to exclude conditions other than PMR. Based on the proposed diagnostic criteria and the exclusion criteria PMR was finally diagnosed by the experienced clinician. All patients met the Bird/Wood criteria.

Seventy-six patients (69 women, 7 men) with a mean (SD) age of 68.7 (7.7) years from Jerusalem, Piestany, Ljubljana, Pavia, and Vienna were enrolled into the response criteria evaluation investigation. These patients needed to have proven PMR according to the diagnostic criteria listed above and to agree with follow up visits after the first visit. The main factor in determining whether patients proceeded to the “response criteria” study was the availability of long term follow up. Thus, centres where follow up was not available did not participate in the PMR response criteria study. No patients with normal values for ESR or CRP were enrolled into this investigation. Corticosteroid treatment was started at the first visit (week 0). Steroid dosage and tapering of the drug as well as prescription of concomitant drugs was at the discretion of the local investigator. Drugs other than corticosteroids and non-steroidal anti-inflammatory drugs (NSAIDs) were not allowed for the treatment of PMR during the observation period.

A questionnaire including clinical measures and laboratory values was developed and approved by a consensus meeting of the participating investigators. To standardise clinical evaluation, muscle tenderness was chosen as the only investigator dependent procedure. Laboratory tests were performed locally according to local standards and quality control regulations. Patients were seen at regular intervals from week 0 up to week 48 (fig 1) and the following parameters, previously chosen to monitor therapeutic efficacy by consensus of the investigators at the beginning of the investigation, were determined at all visits: ESR (mm/1st h), CRP (mg/l), γ globulin (g/l), serum iron level (Fe, μmol/l), pain as measured by a visual analogue scale (VAS pain), physician’s global assessment (PGA) also measured by VAS, and the duration of morning stiffness (MST) in minutes. Muscle tenderness (MT) of the upper arm and the thigh and self reported myalgia were assessed with a four point scale (0 = none; 1 = mild; 2 = moderate; 3 = severe), as was the ability to raise the arms (0 = none, 1 = below shoulder girdle, 2 = up to shoulder girdle, 3 = above shoulder girdle). Presence of headache, swelling of the temporal artery and masseter claudication (yes/no) were registered, as well as the daily corticosteroid dose (mg prednisolone equivalent), the corticosteroid response time (hours), and the daily NSAID dose.

By week 48, 35 patients had been lost to follow up. Therefore, to ensure the evaluation of a critical number of patients (n = 57), week 24 was chosen as an end point for final analysis (fig 1).

To validate the results achieved in the original patient group 24 additional patients (17 female, 7 male, mean age 71 years (52–85), recruited from the Lower Austrian Centre of Rheumatology, were studied, applying the same set of parameters. These patients were recruited between 1999 and 2001 following the same protocol used for the international cohort, but were not enrolled in the PMR diagnostic criteria study. This cohort of patients was evaluated at weeks 0, 4 and at week 33.6 (24.5). No major differences in demographic data were seen compared with the original cohort (data not shown).

Table 1 Investigations stipulated to exclude other conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Investigation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatoid arthritis</td>
<td>Rheumatoid factor, x ray of hands,</td>
<td>x ray, and any other affected joint</td>
</tr>
<tr>
<td>Capsulitis of shoulder</td>
<td>Full examination of shoulder with</td>
<td>clinical history</td>
</tr>
<tr>
<td>Cervical spondylosis</td>
<td>Clinical history, x ray of cervical</td>
<td>spine</td>
</tr>
<tr>
<td>Osteoarthritis of shoulder</td>
<td>x Ray</td>
<td></td>
</tr>
<tr>
<td>Polymyositis</td>
<td>CK, electromyography where clinically indicated</td>
<td></td>
</tr>
<tr>
<td>Thyrotoxicosis</td>
<td>Thyroid function test</td>
<td></td>
</tr>
<tr>
<td>Myopathies</td>
<td>Clinical examination, compatible</td>
<td>electromyography</td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>ANF, DNA binding</td>
<td></td>
</tr>
<tr>
<td>Polyarteritis nodosa</td>
<td>Clinical picture</td>
<td></td>
</tr>
<tr>
<td>Dermatomyositis</td>
<td>Clinical picture, CK</td>
<td></td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td>Serum and urine electrophoresis</td>
<td></td>
</tr>
<tr>
<td>Oesophagus carcinoma</td>
<td>Careful clinical history, chest x ray, occult bloods</td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>Clinical history and clinical signs</td>
<td></td>
</tr>
</tbody>
</table>

CK, creatine kinase; ANF, antinuclear factor.

RESULTS

As was to be expected patients showed a quick and impressive response to the introduction of corticosteroids. The initial dose (mean (SD)) was 24.68 (28.61) mg prednisolone equivalent, which could be tapered to 7.68 (3.61) mg at week 24 (p<0.0001). The corticosteroid response time was evaluated by asking the patient for the onset of improvement after the first corticosteroid dose and amounted to 35.4 (19) hours. As the dose of corticosteroids was at the discretion of the treating doctor the absolute value and the standard deviation of the corticosteroid response
time can be regarded as dependent upon the initial dose chosen by the different investigators.

In the second cohort, the initial (mean (SD)) corticosteroid dose was 27.35 (18.86) mg prednisolone equivalent and 4.23 (3.65) mg at the final control, which was also highly significantly different (p<0.0001).

Changes of disease activity markers
ESR and CRP, VAS pain, PGA, MST (data not shown), myalgia, MT, and the ability to raise the arms (fig 3) all showed highly significant improvement (p<0.0001) at week 24 compared with week 0. Iron levels did not change significantly (table 2).

As measured by an increase of the VAS pain of 10 or more, 11 patients (21%) experienced 15 episodes of relapse, which were accompanied by another increase of CRP levels in eight cases.

Headache (n = 12), swelling of the temporal arteries (n = 7) and masseter claudication (n = 2) were too infrequently reported at the beginning of the investigation and consequently these measures were not entered into the statistical analysis. At week 4 five patients reported headache and three patients were found to have a temporal artery swelling while no patient had masseter claudication. At week 24 only one patient had a headache.

When the Spearman rank correlation was used all parameters that showed significant changes throughout the observation period were also found to be significantly correlated with VAS pain (p<0.0001) and PGA (p<0.0001).

Multiple regression analysis using VAS pain as the dependent variable showed that the changes of ESR (p = 0.08), CRP (p = 0.41), α₂ globulin (p = 0.13), MST (p = 0.1), and MT (p = 0.07) did not correlate with VAS pain, whereas myalgia (p<0.001) and elevation of the upper limbs (p = 0.003) did correlate with VAS Pain. PGA and VAS pain were statistically highly significantly correlated using regression analysis (fig 4) and the Spearman rank correlation, but as PGA constituted the only patient independent parameter it was excluded from the multiple regression analysis.

The mean (SD) overall response rate at week 4 (n = 76) amounted to 52.8 (26)% and at week 24 (n = 57) to 74 (27.8)% which is significantly different (p<0.001). Comparing the individual response rates of the patients evaluated at week 24 with the rates at week 4, 50 patients (88%) were found to show an identical or increasing response, while in seven patients (12%) the response rate had decreased.

Development of response criteria core set
Based on the results described above, we propose a core set of markers of disease activity for monitoring therapeutic responses in PMR. These markers comprise ESR or CRP, representing the acute phase response, and VAS pain, PGA, MST, and the ability to raise the arms, representing the clinical situation.

Regression and correlation analysis showed that every single parameter of the core set significantly influences the individual response rate, with the highest weight for VAS pain (r_s = 0.7419), followed by physician’s assessment (r_s = 0.7273), MST (r_s = 0.6159), CRP levels (r_s = 0.5798).
and the elevation of the upper limbs \((r_s = 0.4555)\) showing the lowest degree of influence on the response rate.

In view of the crucial role of pain during the course of the disease, and because of our findings that demonstrate a close relationship of pain with some other parameters, we regarded changes of pain to be a mandatory marker, while for the other four parameters changes of just three of them reflect changes of disease activity with sufficient accuracy.

In summary, we propose response criteria comprising VAS pain, PGA measured by VAS, MST in minutes, ESR or CRP, and the degree of elevation of upper limbs \((0–3)\), with changes of pain essential in addition to changes in three of the four other parameters (table 3). Percentage changes in the degree of elevation of upper limbs were calculated by 34% (change by 1 point), by 67% (change by 2 points), and 100% (change from zero to three).

Table 4 shows the performance of the response criteria in the European patient group at week 24. The 20%, 50%, 70%, and 90% response rates were calculated either using all five core parameters, or using the proposed response criteria as defined above. Based on the proposed response criteria, in more than the half of the patients (54%, \(n = 31\)) a 90% improvement was shown at week 24.

**Results of the second cohort**

The 24 further patients showed response rates almost identical to those of the larger international cohort (fig 5). All patients but three (88%) showed 50% improvement or more at week 4 (mean 58.3 (21.1)). At the final control (week 33) mean CRP levels had improved by up to 80.5%, elevation of upper limbs by up to 35.8%, morning stiffness was reduced to 69.8%, pain to 74%, and physician’s assessment to 73.6%.

In summary, all patients showed improvement of 50% at least (mean overall response rate 65%), 15/24 (63%) achieved 70% and three patients 90% (13%) improvement. No statistically significant differences could be found between the European and the Austrian patient cohort. When the individual response rates of the patients at the final control visit (week 33 mean) were compared with the rates at week 4, 20 patients had an identical or increasing response, while four patients (17%) showed a decrease of response, which is in line with the relapse rates for the international patient group.

As a consensus an improvement of more than 50% was considered to be clinically meaningful. At week 4 45/76 (59%) patients showed at least 50% improvement of ESR, 68% (n = 52) of CRP, 70% (n = 53) of VAS pain, 68% (n = 52) of PGA, 71% (n = 54) of MST, and 24% (n = 18) of the elevation of upper limbs; and at week 24 74% (n = 42) of the patients showed at least 50% improvement of ESR, 86% (n = 49) of CRP, 82% (n = 47) of VAS pain, 89% (n = 51) of PGA, 95% (n = 54) of MST, and 37% (n = 21) of the elevation of upper limbs.

**DISCUSSION**

As a first result of our study the expected rapid and sustained improvement of PMR after initiation of corticosteroids was proved. Along with the reduction of corticosteroids almost all measures of disease activity, proposed by consensus of the experts, showed significant improvement indicating decrease of inflammatory activity, reduction of pain, and amelioration of functional status of the patients.

A significant change was seen between the response rates at week 4 and at week 28 in the European cohort and a trend for further improvement at the final control for the Austrian patient cohort, respectively. This difference might be due to the diversity of the patients and doctors in the international cohort. The overall response and relapse rates, however, showed no statistical significant differences between the two patient cohorts and are in line with results published previously. The response criteria proposed here may provide a standardised definition of relapse in the future.

**Table 2** Numerical changes of activity parameters in the original patients

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Week 0 (n = 76)</th>
<th>Week 24 (n = 57)</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESR (mm/1st h)</td>
<td>67.37 (30.38)</td>
<td>24.72 (16.97)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CRP (mg/l)</td>
<td>48.59 (47.67)</td>
<td>8.81 (8.72)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Patient assessment (VAS)</td>
<td>70.7 (25.78)</td>
<td>11.91 (20.76)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Doctor assessment (VAS)</td>
<td>63.52 (27.0)</td>
<td>20.76 (10.33)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Morning stiffness (min)</td>
<td>77.24 (75.8)</td>
<td>20.00 (45.95)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Elevation of upper limbs (0–3)</td>
<td>1.83 (0.72)</td>
<td>2.75 (0.54)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Tenderness (0–3)</td>
<td>1.61 (0.98)</td>
<td>0.37 (0.64)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Steroids (mg)</td>
<td>24.68 (28.61)</td>
<td>7.68 (3.61)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

**Figure 4** Regression analysis VAS pain v VAS physician’s global assessment; \(R^2 = 0.954\) in the European patient cohort.

**Table 3** Proposed core set of response criteria

Proposition of a core set of response criteria:
- VAS pain (obligatory)
- CRP (mg/l) or ESR (mm/1st h)
- Morning stiffness (min)
- Elevation of upper limbs (0–3)
- VAS doctor’s global assessment

Three of the last four items must change.

**Table 4** Number (%) of patients achieving 20, 50, 70, and 90% improvement at week 24 in the European patient group using five or four measures of the response criteria core set

<table>
<thead>
<tr>
<th>Week 24 (n = 57)</th>
<th>20%</th>
<th>50%</th>
<th>70%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Parameters</td>
<td>53% (93)</td>
<td>50% (88)</td>
<td>44% (77)</td>
<td>6% (7)</td>
</tr>
<tr>
<td>4 of 5</td>
<td>54% (93)</td>
<td>51% (89)</td>
<td>46% (81)</td>
<td>31% (54)</td>
</tr>
</tbody>
</table>
The measures particularly chosen to be part of the response criteria core set proved their sensitivity to change throughout the observation period up to week 24 (European cohort) or the final control visit of the Austrian patient cohort, respectively.

ESR, CRP, and α2 globulin were chosen to represent the acute phase response and every single parameter showed highly significant improvement. For reasons of practicability CRP and ESR were chosen as part of the response criteria core set with a preference for CRP as ESR may be influenced by other factors to a greater degree, especially in an older population.19 20 Thus, percentage changes of CRP were more pronounced than those of ESR.

Our findings emphasise that pain has the dominant role in a patient’s symptoms. Likewise, it is the principal feature of all diagnostic criteria for PMR published so far.13–15 Therefore, we elected pain status as the central measure for disease activity. Interestingly, our data show that, physician’s assessment and patient’s pain status were highly significantly correlated using three statistical methods. Other parameters, such as myalgia and elevation of upper limbs, were pain dependent, whereas MT did not significantly correlate with spontaneous pain as reported by patients and were pain dependent, whereas MT did not significantly correlate using three statistical methods. Other parameters, such as myalgia and elevation of upper limbs, were pain dependent, whereas MT did not significantly correlate with spontaneous pain as reported by patients and were pain dependent, whereas MT did not significantly correlate using three statistical methods.

Moreover, no relationship between the acute phase reactants or MST and pain status was seen, indicating that pain and the degree of inflammation are independent markers of disease activity.

Thus a core set of response parameters was established consisting of VAS pain as the central measure of the patients’ symptoms, with the acute phase reactants (CRP or ESR) and MST as pain independent variables of disease activity. In addition PGA was included in the core set as the unique functional status the ability to raise the arms, although shown to be slightly pain dependent, was also included instead of MT, which might have resulted in higher response rates but no consideration of functional possibilities. Indeed the inclusion of the ability to raise the arms may lower the sensitivity of the PMR criteria, because this measure might also be abnormal in other disorders, such as osteoarthritis or degenerative disease of the cervical spine, particularly in the population affected by PMR.21 Moreover, the possibility of achieving a 70% response on a four point scale is considerably limited.

Another limitation of the applicability of the response criteria proposed might be given by the fact that PMR may also occur without, or with only a mild, increase of ESR or CRP. Although this may predict less severe disease, in such cases measurement of the acute phase reactants may not reflect disease activity accurately.22

For all those reasons it was decided to insist on the presence of a change of pain intensity, whereas of the other four parameters only three had to change to indicate improvement or deterioration. Moreover, changes of pain intensity contributed most to the response rates. Although it is obvious that neither diagnostic criteria nor response criteria can cover every single patient affected by the disease, we demonstrate in two different cohorts of patients with PMR that the response criteria core set proposed is applicable for a high percentage of patients with PMR. In more than half of the patients a 70% response could be achieved at week 24 in the European cohort and in the Austrian cohort respectively, which is in line with the excellent prognosis of PMR often described. Achieving 70% and 90% improvement comes near to remission, although definition of remission is still lacking for PMR. Therefore the calculation of response rates using changes of pain and two of the other four parameters was ultimately not applied, although it was considered. Using this model would have resulted in a 100% response in almost the half of the patients. It should be the matter of further investigations whether the response criteria proposed here might also be applied to define remission in PMR.

The response rates of the second cohort for corticosteroid dose reduction and course of the single parameters as well as the global response rates were almost identical.

In summary, the PMR response criteria core set presented in this paper are a promising new tool for monitoring disease activity and achieving successful treatment in patients with PMR. Moreover, these criteria are easily applicable in daily practice as well as in clinical trials and may therefore provide an opportunity to compare different patient cohorts and other therapeutic strategies with the “gold standard” of corticosteroids. We would like to encourage rheumatologists to validate and assess the proposed response criteria core set in their particular clinical setting. We expect that better monitoring of treatment responses in PMR will improve clinical decision-making and, ultimately, improve patient care.

ACKNOWLEDGEMENTS
The authors thank the other members of the European PMR collaborating group for their support and important contributions in preparing and conducting this work: R Caporali, Pavia; H Haentzschel, Leipzig, N Mistumiene, Kaunas; S Pai, Tartu; C Pease, Leeds, UK; B Rozman, Ljubljana; A Tuchynova, Piestany; U Zimmermann-Gorska, Poznan.

Authors’ affiliations
B F Leib, I Andel, J Sautner, Lower Austrian Centre for Rheumatology, Humanis Klinikum Lower Austria, Stockerau, Austria
H A Bird, University of Leeds, Leeds, UK
G Nesher, M Sonnenblick, Shaare Zedek Medical Centre, Jerusalem, Israel
W Hueber, Centre for Rheumatic Diseases, Lainz Hospital, Vienna, Austria
D Logar, Department of Rheumatology, Clinical Centre, Ljubljana, Slovenia
C M Monteucco, IRCCS, Policlinico San Matteo, Pavia, Italy
J Rovensky, Research Institute of Rheumatic Diseases, Piestany, Slovak Republic
REFERENCES