

An evaluation of criteria for polymyalgia rheumatica

H. A. BIRD,¹ W. ESSELINCKX,¹ A. ST. J. DIXON,¹ A. G. MOWAT,²
AND P. H. N. WOOD³

From the South Wales and West Country Rheumatology Club and Royal National Hospital for Rheumatic Diseases, Bath¹; Rheumatism Unit, Nuffield Orthopaedic Centre, Oxford²; and Arthritis and Rheumatism Council Epidemiology Research Unit, University of Manchester Medical School, Manchester³

SUMMARY There has been little basis on which to standardise a diagnosis of polymyalgia rheumatica (PMR), and so 11 rheumatology units in the south and west of Great Britain have collaborated in a study to evaluate possible criteria. Symptoms and laboratory findings claimed to be of diagnostic value in PMR were included in an analysis of the features of 236 patients considered to have unequivocal PMR and 70 patients thought to have possible PMR. The results were compared with similar information from 253 patients with conditions that mimic PMR and from 201 consecutive new presentations to outpatients. The 7 most valuable criteria for differentiation were bilateral shoulder pain or stiffness, onset of illness of less than 2 weeks' duration, initial ESR greater than 40 mm/h, duration of morning stiffness exceeding 1 hour, age 65 years or more, depression and/or weight loss, and bilateral tenderness in the upper arms. We suggest that a patient might be regarded as having probable PMR if any 3 or more of these criteria are fulfilled, or if at least 1 criterion coexists with a clinical or pathological abnormality of the temporal artery. A standardised therapeutic test with prednisolone has value in making the diagnosis of PMR more certain.

The term polymyalgia rheumatica (PMR) was introduced by Barber in 1957. However, it is probable that the descriptions of senile rheumatic gout (Bruce, 1888), mylagic syndrome in the elderly (Kersley, 1951), anarthritic rheumatoid syndrome in the elderly (Bagratuni, 1953, 1963), and pseudo-polyarthrite rhizomélique (Forestier and Certoncini, 1953) all referred to the same condition. These designations reflect the variety of ways in which PMR may present, and published series have each given their own emphasis to different features. As with so many rheumatic disorders the fundamental problem is the lack of a single defining feature or hallmark, such as a specific diagnostic test. Attempts have been made to develop criteria for identification in situations of comparable difficulty, such as with rheumatoid arthritis (American Rheumatism Association, 1957) and systemic lupus erythematosus (Arthritis Foundation, 1971), but analogous efforts for PMR appear to have been neglected. Diagnostic criteria cannot be used in such a way that a disease can be said not to be present if the manifestations fail

to fulfill the criteria. However, criteria are required to ensure comparability of diagnosis when different authors report features such as the frequency of complications, associations with other conditions, response to various treatments, or prevalence in different communities.

In the case of PMR the situation is confused further by the overlap with temporal (giant cell) arteritis; the features encountered in the prodromal phase of the latter condition are identical with those of PMR. Temporal arteritis was described first by Hutchinson in 1890, but many regard the two conditions as part of the same disease spectrum (for example, Dixon *et al.*, 1966; Fauchald *et al.*, 1972; Mowat and Hazleman, 1974).

For our understanding to increase it is obviously necessary for reproducible means of identifying conditions like PMR to be developed, and so 11 rheumatology centres in the south and west of Great Britain have collaborated in an evaluation of possible diagnostic criteria.

Patients and methods

Sixteen physicians at 11 centres submitted data on patients they considered to have either unequivocal

Accepted for publication 2 October 1978
Correspondence to Dr P. H. N. Wood, ARC Epidemiology Research Unit, Stopford Building (University of Manchester), Oxford Road, Manchester M13 9PT.

or possible PMR. Details of history, examination, and laboratory investigations were transferred to a mark-sense computer coding form by H.A.B. We endeavoured to include all features that have been claimed to have diagnostic value for PMR, and these are enumerated in Table 2. One characteristic requires clarification; what we have termed 'onset of illness of <2 weeks duration' refers not to the stage in their illness at which patients were seen, but to the time taken for symptoms to reach their full-blown picture from the patient's point of view, an aspect elicited in the history. We also recorded standard demographic data, clinical features, changes in the erythrocyte sedimentation rate (ESR), and the speed and nature of response to treatment. We assessed the latter in 3 ways simultaneously: (1) rapidity of at least 50% overall clinical improvement according to the patient's own retrospective appreciation; (2) progress in pain, stiffness, and general wellbeing recorded by the patient daily on separate visual analogue scales (VAS) for 14 days, the VAS being returned by post every day after completion; and (3) the change in ESR at the end of a 14-day course of prednisolone, 10 mg/day. However, although response to corticosteroid therapy is widely believed to have value as a discriminator for PMR, one would prefer to make a provisional diagnosis before carrying out a therapeutic test; we therefore excluded this characteristic from most of our analyses.

The basic study design is indicated in Table 1. Identical information was collected from 146 patients with unequivocal PMR seen at 2 of the centres, Bath and Oxford, and from 253 patients with conditions that often mimic PMR. These diseases that may give rise to diagnostic confusion included rheumatoid arthritis, local conditions of the shoulder, osteoarthritis, polymyositis, cervical disc lesions, metabolic disorders (e.g., thyrotoxicosis), myopathies, systemic lupus erythematosus, polyarteritis nodosa, dermatomyositis, multiple myeloma, carcinomatosis, and Parkinson's disease. This basic comparison with a homogeneous group of patients with PMR allowed features to be tabulated in order of discriminatory

performance (Table 2), and from this list we selected those that performed best.

Table 2 Clinical and laboratory characteristics of 146 patients with unequivocal PMR (compared with features of 253 patients with conditions giving rise to diagnostic confusion; features ranked by magnitude of relative values—the sum of sensitivity and specificity)

Characteristic*	Sensitivity (%)	Specificity (%)
Shoulder pain and/or stiffness bilaterally ¹	86	68
Onset of illness of <2 weeks duration ²	88	63
Initial ESR ≥ 40 mm/h (but tested separately by 10 mm/h increments)	74	74
Morning stiffness duration > 1 h ³	80	61
Age > 65 yr (but tested separately by quinquennial increments)	70	69
Depression and/or loss of weight	58	81
Upper arm tenderness bilaterally ⁴	36	96
Loss of appetite ⁵	37	95
Exacerbation of symptoms by exertion—absence of ⁶	94	27
Swelling of knee—absence of ⁷	92	23
Limitation of motion of shoulder bilaterally ⁸ (external rotation <90°)	27	85
Leucocyte count > 6000/cu mm (6 × 10 ⁹ /l)	91	21
Temporal artery abnormality unilaterally or bilaterally (including pulsation, tenderness, beading, and bruit) ⁹	12	100
Headache or pain (including frontal, temporal, occipital, and facial)	24	87
Eye symptoms (including loss of vision, diplopia, and orbital ache)—absence of	97	14
Haematocrit < 40%	56	54
Haemoglobin < 11 g/dl	16	93
Fever (history of; measured temperature less satisfactory as discriminator)	10	96
Ischemic pain—at any site (including gastrointestinal tract)—angina	11	95
Alkaline phosphatase ≥ 13 King-Armstrong units	9	97
Ischemic pain—in jaw—intermittent claudication (of calf)—absence of	29	76
Rheumatoid factor test negative	4	99
Limitation of motion of hip bilaterally (abduction < 45°)	100	1
	89	11
	3	93

*Characteristics are identified in two ways; either by specification of the precise features considered, when alternative formulations (e.g. unilateral as opposed to bilateral) have also been examined and found to perform less satisfactorily, or by enumeration of different attributes that were recorded separately (noted by inclusion statements, some of which are given in numbered footnotes), when the results were examined for the presence of each attribute individually as well as in combination. Where appropriate characteristics have been examined as negative features—i.e. absence of the feature being regarded as the characteristic under test. The frequency of missing observations, mainly in the comparison series, precluded satisfactory tabulation of the following features: fall in ESR, erythrocyte count, proportions of polymorphonuclear and eosinophil leucocytes, serum glutamic pyruvate transaminase (SGPT), plasma proteins, proportion of albumen, and biopsy material.

¹ Pain and/or stiffness recorded for following sites: neck, shoulder, upper arm, buttock, and thigh, and, as negative features, for hand, back, knee, and calf; data cited refer to best discriminator.

² Refers to time taken for symptoms to reach their full-blown picture.

³ Full data recorded included whether stiffness worse by night (i.e. morning stiffness) or whether it was gelling (i.e. rest stiffness), and its duration.

⁴ Better discriminator than bilateral thigh tenderness.

Table 1 Groups of patients studied for evaluation of criteria for polymyalgia rheumatica

Clinical and laboratory features compared in patients with:	
—unequivocal PMR from 2 centres	n = 146
—conditions giving rise to diagnostic confusion	n = 253
Validation by comparison in additional series of patients with:	
—unequivocal PMR from 9 other centres	n = 90
—new consecutive outpatient presentations	n = 201
Additional evaluation by prospective study of patients with:	
—possible PMR	n = 70

⁵ Loss of appetite has been shown below the line because, although the sum of its sensitivity and specificity was similar to that for valuable characteristics, it was assumed that this feature replicated information conveyed by weight loss, and the latter was a more objective indicator.

⁶ Full data recorded included exacerbation by exertion, slightest motion, and pressure (such as in buttocks when sitting down for a while).

⁷ Swelling included effusion and/or synovial thickening and was recorded in sternoclavicular joint, knee, leg (oedema), and elsewhere; best discriminator shown.

⁸ Contrast findings in the hip (last entry in Table).

⁹ Other arteries examined similarly included carotid, subclavian, axillary, brachial, radial, femoral, and popliteal; best discriminator shown.

The provisional criteria derived in this way were then validated in a second comparative study between a further 90 patients with unequivocal PMR from 9 other centres, and 201 consecutive new patients presenting at the rheumatology outpatient department in Bath. A further validation was also initiated by studying the performance of the criteria in 70 patients with possible PMR. Prospective follow-up should resolve the diagnosis in these cases, allowing evaluation of the criteria in less definite circumstances.

ASSESSMENT OF DISCRIMINATORY PERFORMANCE

In each comparative study the sensitivity and specificity of every individual feature was calculated by computer. Sensitivity is the proportion of patients with the disease who are positive for the feature, the true positives, when related to all individuals with the disease; and specificity is the proportion of patients without the condition who are negative for the feature, the true negatives, when related to all individuals without the disease. To be useful a diagnostic criterion should be both sensitive and specific. Perfection is rare, however, and we have to accept compromises that are less than ideal. The simplest way of making these choices is by consideration of the relative value of individual criteria. This is the sum of sensitivity and specificity, and, as both of these are expressed as percentages, when added together they can range from 0 to 200. We think relative value is easier to comprehend than the Youden Index (Blumberg, 1957). The latter, designed to indicate the optimal screening level for a feature, has a similarly empirical basis but tries to express results on a percentage scale by subtracting 100 from the sum of sensitivity and specificity.

Ideally the information content of any individual criterion should not appreciably replicate that conveyed by other criteria; failure to observe this caution leads to the attachment of undue weight to the features represented more than once. For example, to include pain or stiffness experienced at each of a number of proximal sites is really a duplication of information. Thus, in this example,

although we considered each site individually, the neck, shoulder, upper arm, buttock, thigh, and so on, we found that confining attention to the shoulder yielded slightly better discrimination than taking pain or stiffness at each of the proximal sites alone or in combination; thereafter we treated this as a single characteristic, although the relative values of the alternatives were usually of a similar, even if slightly lower, order of magnitude. Similarly, although we examined many arteries for tenderness, bruits, and absence of pulsation, we found most satisfactory discrimination when we considered only abnormalities in the temporal artery.

Necessary though this process of preliminary sifting may be, there is a degree of artificiality in considering individual features in isolation. This arises because conclusions in clinical practice are usually based on the simultaneous presence of different characteristics. The same analytical procedures were therefore then applied to combinations of those criteria that had the best discriminatory performance. The sensitivity and specificity of criteria taken in conjunction serve to indicate the probabilities of patients suffering from PMR or from some condition other than PMR in relation to the number of characteristics present.

Results

The performance of individual features in the basic comparative study is shown in Table 2, where the characteristics have been ranked by the magnitude of the sums of sensitivity and specificity. There is an obvious separation between the characteristics. This distinction is emphasised by a horizontal rule, features above the line having the greatest value for discrimination.

Fig. 1 shows the accumulated proportions of patients fulfilling various numbers of criteria for the 7 characteristics with the highest relative values as diagnostic criteria. Thus in the upper part of Fig. 1 the first column indicates that all patients with PMR fulfilled at least 1 criterion, whereas only 97% fulfilled 2 or more, indicated in the second column. Specificities are shown similarly in the lower part of Fig. 1, and the relative value for any given number of criteria is reflected by the sum of the areas in the pair of columns situated one above the other. The optimal relative value is the highest estimate, and, as recorded in a footnote to Fig. 1, this was found for 2 levels, indicated by *x* and *y* and relating to 3 or more and to 4 or more criteria respectively. There is obviously little to choose between these values, and in such a situation the differences in sensitivity are critical. Thus if we took level *y* we should miss every fifth case of PMR, as the sensitivity at this level was

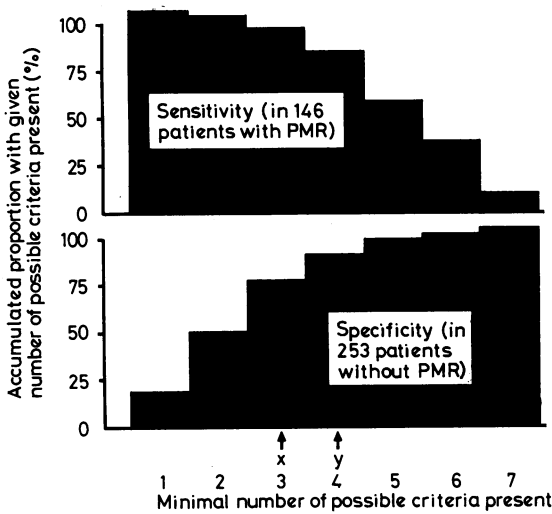


Fig. 1 Accumulated proportions of patients with PMR with various numbers of possible diagnostic criteria (proportions are accumulated because any one value indicates the proportion of patients fulfilling at least that number of possible criteria). Optimal relative values of 172 and 171 respectively are indicated by the arrows marked x and y.

only 78%. Level x is obviously a better choice, 3 or more criteria, because discrimination is just as good and yet only 1 PMR in 12 is missed (92% sensitivity).

Assessment of the discriminatory performance of individual characteristics was repeated in the validation series, and the results are summarised in Table 3. Despite the fact that there was less homogeneity, both of individuals with PMR and of the patients without PMR with whom they were being compared, 5 of the 7 characteristics had similar value. The only exceptional results were in relation to morning stiffness, depression or weight loss, and rheumatoid factor. The latter had a higher specificity as a negative feature so that its discriminatory value improved, in contrast to all the other less valuable

Table 4 Application of proposed diagnostic criteria for PMR

	Basic study	Validation study	Possible PMR
Sizes of samples:			
—of comparison group	253	210	201
—of case material	146	90	70
Probable PMR identified on basis of:			
—fulfilling 3+ criteria	134	72	47
—fulfilling <3 criteria but with temporal artery abnormality	0	2	1
Discriminatory performance of probable PMR:			
—overall sensitivity (both above categories)	92%	82%	69%
—specificity (based on 3+ criteria)	80%	75%	75%
Proportion of probable cases with corticosteroid response			
	96%	97%	74%

characteristics shown in Table 2 which in general showed marginally less discrimination on validation.

Given this degree of stability in the data we were able to validate the performance of possible criteria taken in combination, and the results of these analyses are shown in Table 4. If probable PMR was identified on the basis of fulfilment of 3 or more criteria, a sensitivity of 80% was obtained in the validation series and of 67% in the cases of possible PMR. The association with temporal artery abnormalities was also considered in those not fulfilling 3 criteria, and, although this added very few cases, the overall results are shown in Table 4. Finally, the possible confirmatory value of a therapeutic response to corticosteroids was examined in those with probable PMR, and these results are also shown.

Discussion

The 7 characteristics giving best discrimination highlighted features of PMR that were much what might have been expected, indicating elderly patients with an illness of which the major features had manifested themselves within 2 weeks and in which proximal pain or stiffness and a high ESR were

Table 3 Validation of discriminating characteristics for PMR (replication of analyses in second comparative series—see Table 1)

Characteristic	Basic study		Validation study	
	Sensitivity (%)	Relative value	Sensitivity (%)	Relative value
Shoulder pain and/or stiffness bilaterally	86	155	83	158
Onset of illness of <2 weeks duration*	88	151	79	134
Initial ESR ≥ 40 mm/h	74	149	87	152
Morning stiffness duration > 1 h	80	141	51	87
Age > 65 years	70	139	73	141
Depression and/or loss of weight	58	139	29	109
Upper arm tenderness bilaterally	36	132	34	130

*See footnote 2 to Table 2.

noteworthy. Our patients also showed the seasonal clustering described by Mowat and Hazleman (1974), but no clear-cut familial, neighbourhood, or geographical patterns emerged in the area of south-west Great Britain from which our patients were drawn.

Our patients also showed many other features that have been associated with PMR, such as pyrexia and a prodromal virus illness (Table 2). It is easy to appreciate why features like these have been suggested as having diagnostic value. In most instances they had a high specificity rather than being particularly sensitive. Temporal artery abnormalities epitomise the situation; though they were encountered almost exclusively in PMR, only 1 patient in 10 had this feature. The distinction is between what is highly characteristic and what serves to discriminate from other conditions. Except in the minority who manifest the characteristic, criteria with low sensitivities generally contribute little to arriving at a diagnosis: the diagnostic value of subcutaneous nodules in rheumatoid arthritis is similar. The performance of haematocrit was exceptional in that sensitivity and specificity were of similar magnitude, but overall this characteristic had little discriminatory value. Of the other less valuable characteristics it is worth noting that the apparently high sensitivities of some are rather artificial, being a reflection of the fact that these attributes were treated as negative features and that their indices therefore are really the reverse of those obtained with positive attributes.

The general stability of the results on validation in different groups of patients is certainly encouraging. However, the poorer performance of morning stiffness and depression or loss of weight calls for further study. Nevertheless the unsatisfactory aspect of these latter characteristics is of less importance when criteria are looked at in combination. The simple model proposed, that presence of 3 or more of the specified characteristics is sufficient to permit identification of probable PMR, held up well on validation; the fall in relative value from 171 to 157 when comparisons were extended to less selected case material is in fact unusually encouraging. It must be acknowledged, though, that we have used only simple and pragmatic analytical approaches so far, and that our model is therefore only provisional. We are pursuing more sophisticated methods to see if we can maximise discriminatory performance. One approach is not to match a patient against a set of characteristics but instead, as Wood (1978) has suggested, to apply a series of discriminators in sequence. Our use of steroid response as a confirmatory measure is a simple example of this method, and one that performed well. In the meantime the stimulus for other readers is perhaps less to try to propound alternative models, but more to

gather comparable data, so that the basis of this collaborative effort to identify PMR more reproducibly might be extended. The performance of our model in patients with possible PMR is also encouraging, but this requires further validation by follow-ups to observe what became apparent with the passage of time.

At the outset we noted the conceptual problem associated with the overlap in features between PMR and temporal arteritis. Our work does nothing to clarify this situation, but it does provide the means for pursuing studies in this area. Thus observation of diagnostic characteristics in patients with PMR but without temporal arteritis and in patients with apparently pure temporal arteritis could help to resolve some of the difficulties.

Three minor difficulties require acknowledgement. First, although we tried to be comprehensive we inevitably failed to study certain features not recorded in the literature. For example, Dr M. Gumpel has stressed to us the patient's ability to roll over in bed at night and whether he or she has to wake up for this purpose, and perhaps even to get his spouse to assist with a shove. Secondly, some readers may be unhappy at the attempt to record the time during which the illness unfolded, particularly because patients may have a history extending over many months before they get referred to a rheumatologist. However, we did not experience difficulty in eliciting this valuable criterion. (The delayed referral problem is something distinct, to a great extent reflecting the accessibility of a consultant opinion and the success of efforts to alert general practitioners to the frequency and characteristics of PMR.) Thirdly, objective assessment of the therapeutic response to corticosteroids is unsatisfactory, and further study is required so that, among other things, dosage can be standardised. Dixon (1978) has proposed a single-blind diary-monitored design for the test that helps in this regard.

We thank the following physician members of the South Wales and West Country Rheumatology Club for their help and co-operation: Drs P. A. Bacon and J. A. Cosh (Bath), M. I. V. Jayson (Bristol), J. C. Harding (Bournemouth), J. D. Jessop, K. N. Lloyd, and G. Nuki (Cardiff), G. H. Hall (Exeter), J. L. Milligan (Poole), A. K. Tyler (Portsmouth), M. I. D. Cawley (Southampton), P. N. Knight (Swindon), D. B. Yates (Taunton), and R. K. Jacoby (Torbay). We are grateful to David Hewitt, Denise Gorman, and Trevor Benn for valuable help in document preparation and data handling, and to Miss N. Russell, Miss E. L. Lupton, and Mrs J. Battersby for secretarial assistance. This work was carried out with support from the Arthritis and Rheumatism Council.

References

- American Rheumatism Association (1957). Proposed diagnostic criteria for rheumatoid arthritis. *Annals of the Rheumatic Diseases*, 16, 118-125.

- Arthritis Foundation (1971). Preliminary criteria for the classification of systemic lupus erythematosus. *Bulletin on the Rheumatic Diseases*, **21**, 643–648.
- Bagratuni, L. (1953). A rheumatoid syndrome occurring in the elderly. *Annals of the Rheumatic Diseases*, **12**, 98–104.
- Bagratuni, L. (1963). Prognosis in the anarthritic rheumatoid syndrome. *British Medical Journal*, **1**, 513–518.
- Blumberg, M. S. (1957). Evaluating health screening procedures. *Operations Research*, **5**, 351–360.
- Bruce, W. (1888). Senile rheumatoid gout. *British Medical Journal*, **2**, 811–813.
- Dixon, A. St. J., Beardwell, C., Kay, A., Wanka, J., and Wong, Y. T. (1966). Polymyalgia rheumatica and temporal arteritis. *Annals of the Rheumatic Diseases*, **25**, 203–208.
- Dixon, A. St. J. (1978). The diagnosis of polymyalgia rheumatica. *Reports on Rheumatic Diseases*, Collected Reports 1959–1977, pp. 61–63. Edited by C. Hawkins and H. L. F. Currey. Arthritis and Rheumatism Council: London.
- Fauchald, P., Rygvold, O., and Øystese, B. (1972). Temporal arteritis and polymyalgia rheumatica. Clinical and biopsy findings. *Annals of Internal Medicine*, **77**, 845–852.
- Forestier, J., and Certoncini, A. (1953). Pseudo-polyarthrite rhizomélique. *Revue du Rhumatisme et des maladies ostéo-articulaires*, **20**, 854.
- Kersley, G. D. (1951). A myalgic syndrome in the elderly. Proceedings of the Second Congress of European Rheumatology, Barcelona. *Revista Española de Reumatología*
- Mowat, A. G., and Hazelman, B. L. (1974). Polymyalgia rheumatica—a clinical study with particular reference to arterial disease. *Journal of Rheumatology*, **1**, 190–202.
- Wood, P. H. N. (1978). Nomenclature and classification in rheumatology. *Copeman's Textbook of the Rheumatic Diseases*, 5th edn., Chap. 2, pp. 14–24. Edited by J. T. Scott. Churchill Livingstone: Edinburgh.