Collection of data terminated at the end of the year 2000. Although, as this abstract is prepared, some data sets are still awaited, it is likely that analysis will be based on data from 312 patients judged to have polymyalgia rheumatica recruited from 13 centres.

The vast majority of these patients, judged by a consensus of experts to have polymyalgia rheumatica, have provided a database for comparison of the various diagnostic criteria sets.

On the basis of this, optimum European diagnostic criteria will be proposed. Investigators are encouraged to use this in their research on polymyalgia rheumatica to further substantiate their reliability. A handout listing the new criteria will be available at the time of the presentation.

REFERENCES

- 1 Bird HA, *et al.* *Ann Rheum Dis.* 1979;**38**:434–9
- 2 Jones JG, Hazleman BL. *Ann Rheum Dis.* 1981;**40**:1–5
- 3 Nobunaga M, *et al.* *Jpn J Med.* 1989;**28**:452–6

SP0054 POLYMYALGIA RHEUMATICA REMISSION CRITERIA

B Leeb. *Lower Austrian Center for Rheumatology, Stockerau Hospital, Stockerau, Austria* Introduction

10.1136/annrheumdis-2001.13

Polymyalgia rheumatica (PMR) is not a benign disease. Additional morbidity may arise from prolonged administration of high doses of corticosteroids. No stringent response criteria have been established yet, however, they obviously seem to be of great importance for the quality of patient care and to compare alternative treatment modalities against the golden standard of steroid treatment as steroid sparing medications would be most desirable in PMR.

Patients and methods Patients from several European countries were enrolled into a study to validate existing diagnostic criteria for PMR as well as into the remission criteria study. Serological findings (ESR, CRP, haemoglobin, serum iron and alpha2-globulin) and clinical features (VAS pain, VAS physicians assessment, morning stiffness, elevation of upper limbs, myalgia, tenderness on pressure of upper/lower limbs, response time after initiation of steroids, head aches, swelling or tenderness of the temporal artery and masseter muscle claudication) were chosen for the characterisation of disease activity. After the initial examination patients were seen at week 1, 4, 8, 16, 24 and 32, study end-point was at week 48. Starting dose and reduction of corticosteroids as well as co-medication was at the discretion of the local investigator. Patients found to have a disease other than PMR will be excluded from final analysis in the response criteria study.

Results As expected most of the patients responded well and quickly to steroid treatment. The results concerning the single response parameters will be given on date. Consisting of those criteria performing best a response score should be proposed. The efficacy of treatment may thus easily be reflected by a single figure.

SP0055 POLYMYALGIA RHEUMATICA NEW TREATMENT (NEW STEROID SPARING AGENTS)

GG Hunder. *Rheumatology Department, Mayo Clinic, Rochester, USA*

10.1136/annrheumdis-2001.14

Polymyalgia rheumatica (PMR) responds well to glucocorticoids such as prednisone when given in small to medium-sized doses. But because the condition tends to persist for months or even years, adverse effects develop in most patients (Gabriel, *et al.* *Arthritis Rheum* 1997;**40**:1873). A variety of steroid sparing drugs have been tried. One NSAID, Tenidap, showed steroid sparing efficacy but also relatively frequent adverse effects. The new Cox-2 inhibitor drugs have not been studied adequately in PMR to determine if they have a role in certain cases.

Experience with cytotoxic drugs has been mainly in cases of giant cell arteritis (GCA), some of whom also had PMR. Even in GCA, cytotoxic investigations have been limited. Because of the close link between GCA and PMR it would be presumed that drugs useful for GCA would also be effective in PMR. Methotrexate has been studied most. Two recent randomised, double-blind, placebo-controlled studies using methotrexate as a steroid sparing agent are the best work on this question to date, but gave different results. In one (Jover, *et al.* *Ann Intern Med* 2001;**134**:106), 33 patients completed the trial. 18 patients were given prednisone plus placebo, and 15 were given daily prednisone plus 10 mg methotrexate per week. All were followed for 24 months. The group receiving methotrexate with prednisone experienced fewer relapses ($p = 0.02$) and used significantly less prednisone ($p = 0.004$) over the study period. The second study (Hoffman, *et al.* In press) was also a prospective randomised double-blind study. 98 patients were enrolled. 51 received methotrexate (mean dose 15 mg per wk) plus prednisone, and 47 received placebo and prednisone. The results showed that the methotrexate group was not different from the placebo group regarding outcome and prednisone dose. The reasons for the differences between these studies is not completely clear but will be discussed. Thus, at this time there are no steroid sparing drugs that can be relied on for PMR. At this time in the practice setting we would recommend a trial of methotrexate only in patients who at very high risk of developing steroid side effects. Additional investigations with this drug and others is needed.

Clinical epidemiology in practice – Thursday 14 June, 16.00-17.30/Meeting Hall IV

SP0056 MEASUREMENT ISSUES

A Tennant. *Rheumatology and Rehabilitation Research Unit, University of Leeds, Leeds, UK*

10.1136/annrheumdis-2001.15

In a recent European survey of outcome measures for rheumatoid arthritis over 60 different instruments were found to be in use. For low back pain, over 100 different instruments were found to be in use. If, for example, measuring disability is considered important, what are the characteristics of a good measure of disability? How do we choose between competing instruments?

Recent advances in our understanding of outcome measurement have identified a number of fundamental criteria that should be met, prior to more established tests of reliability, validity and responsiveness. The two most important characteristics are unidimensionality and absence of Differential Item Functioning (DIF).

The ICIDH helps us to think conceptually about outcomes, and contributes to measurement by identifying unidimensional

constructs. A new mathematical approach (Rasch analysis) then allows for formal tests of unidimensionality in ordinal scales. Questions, or tasks which are summated in some fashion to give an overall score must be unidimensional. Furthermore, at the same level of ability (for example in walking), the probability of a patient doing a task must be the same irrespective of whether they are male or female, young or old, underlying pathology (cross-diagnostic validity) or country of use (cross-cultural validity) – the absence of DIF.

Once these fundamental requirements of measurement are met, tests of reliability, validity and responsiveness may proceed. In this way, those working in the field of arthritis can be confident that evaluation of interventions are supported by high quality measurement.

SP0059 CROSS-CULTURAL VALIDATION OF PERCEIVED HEALTH STATUS MEASURES

F Guillemin. *School of Public Health, Faculty of Medicine, Vandoeuvre-Les-Nancy, France*

10.1136/annrheumdis-2001.16

Since difference in country, language or culture make people vary in disease expression, in perception of health problems and quality of life, and in their use of health care systems, there is a need for measures tailored to their particular need. It is also increasingly necessary to have standardised instruments for international use in research and practice. To resolve this paradox, the process of cross-cultural validation will particularly focus on the achievement of equivalence in concept measured across culture.

Perceived health status measures are complex scales in the form of interview or self-report questionnaires. Their validation shall follow the standards of psychometric validity (content, construct, reliability and sensitivity to change). The content validity is particularly challenging the cross-cultural validation. Among options to conduct validation of the content, it is possible to achieve it by:

- building an instrument concomitantly in several countries. It allows to work together in the generation and selection of items, as has been done with disease-specific (QUALEFFO, Womac), or generic instruments (WHOQOL).
- developing a standard methodology to cross-culturally adapt an established instrument offering guarantee of equivalence between source and target. The participation of the developers to the process is highly contributive to clarify concept explored in the original language

Researchers are most frequently facing the situation where the original version has been developed and validated in a foreign language. Thus the second option will be particularly useful.

Simple translation/back-translation is far from meeting basic requirements for equivalence in content. Perfect bilingual experts are scarce, if any.

Guidelines are proposed to ensure cross-cultural content equivalence. They include:

- several translations, with appropriate translators
- expert committee review, including various experts (health professionals, patients) and having specified tasks
- several back-translations, with standardised rules
- testing with patients to probe understandability

The place and role of back-translation has been debated, since its contribution to the process may be overestimated. It is as fallible as translations, and not any (back) translator would be capable of assessing proximity between versions.

From qualitative research, it is now clear that a thoughtful committee working on several translations achieved by adequate translators may well achieve the goal.

One aspect of the cross-cultural adaptation is the potential bias introduced when adapting answer modalities. The scalability of the adapted instrument can be assessed using the Rasch modelling statistic.

Other aspects of validation of the cross-culturally adapted measure should be clearly considered separately. As far as the population in the target culture may differ in many aspects compared to the original targeted population, especially in many characteristics relating to culture, a standard procedure of validation should be conducted:

- construct validity will consider how items and dimensions are clearly exploring separate and complementary facets of the concept, i.e. health status, and the coherence of scales (factor structure and internal consistency)
- convergent validity will be documented by the relation to variables exploring a similar construct,
- discriminant validity will be documented by checking how groups with known difference will be recognised as different by the instrument
- reliability will be documented by the ability of the measure to produce the same value when exploring the same phenomenon
- sensitivity to change will be its power to detect intra individual changes over time.

These steps are mandatory requirements to document the equivalence of the measurement, at the same degree of reliability in different cultural settings. Because it is much shorter process than developing a new instrument in each and every country, and since many instruments are already available, the cross-cultural adaptation and validation of existing perceived health status and quality of life measures is highly recommended.

State of the art lecture – Friday 15 June, 10.00-10.30/Congress Hall

SP0060 DISEASE MODIFYING DRUGS IN THE TREATMENT OF RHEUMATOID ARTHRITIS

M Weinblatt. *Rheumatology, Brigham and Women's Hospital, Boston, USA*

10.1136/annrheumdis-2001.17

The last several years have been very exciting for the therapy of rheumatoid arthritis (RA). One major advance has been the acceptance of a core set of disease activity measurements and definition of improvement. The American College of Rheumatology (ACR) core set of disease measurements includes tender and swollen joint count, physician and patient assessment of disease activity, patient assessment of pain, physical function and acute phase reactants (Arthritis Rheum. 36:729,1993). The ACR definition of response requires a 20% improvement in both tender and swollen joint count and a 20% improvement in three of the five remaining ACR core set measures (Arthritis Rheum.38:727,1995). Another disease activity response rate developed by Dutch investigators is called the DAS (Disease