EVALUATION OF A DUTCH VERSION OF THE AIMS2 FOR PATIENTS WITH RHEUMATOID ARTHRITIS

R. P. RIEMSMA,* E. TAAL,* J. J. RASKER,*† P. M. HOUTMAN,‡ H. C. VAN PAASSEN§ and O. WIEGMAN*

*Department of Psychology, University of Twente, Enschede, †Department of Rheumatology, Medisch Spectrum Twente, Enschede, ‡Department of Rheumatology, Medisch Centrum Leeuwarden and §St Francisca Gasthuis, Rotterdam, The Netherlands

SUMMARY

DUTCH-AIMS2, a Dutch version of AIMS2 and successor to DUTCH-AIMS, is an instrument to assess health status among patients with rheumatic diseases. It provides measurements of 12 areas of health status on scales for health status proper, satisfaction, attribution and arthritis impact. We assessed the reliability of its scales in terms of internal consistency and their validity according to both internal standards and external standards. Correctly completed questionnaires were returned by 231 RA patients and 131 controls. Internal consistency coefficients for the health status scales ranged from 0.66 and 0.89, but most exceeded 0.80. Within-scale factor analyses produced single factors in all composite health status scales for both patients and controls, with only two exceptions. Factor analysis also identified a physical, social and psychological dimension among 11 areas of health. External validity was established by strong correlations between DUTCH-AIMS2 health status scales and functional class, laboratory parameters, and self-assessments of fatigue, loneliness, pain, functional disability and social support. DUTCH-AIMS2 is acceptably reliable and valid for use in a variety of settings.

KEY WORDS: DUTCH-AIMS2, Evaluation, Rheumatoid arthritis.

The Arthritis Impact Measurement Scales (AIMS) formed one of the first instruments to be designed specifically to assess health status among patients with rheumatic diseases [1]. Its measurement properties are good and it has come to be widely accepted in the English-speaking countries for a variety of uses [2-4]. Its Dutch translation, the DUTCH-AIMS, was the first such instrument in the Dutch language. This too has been shown to be acceptably reliable and valid [5, 6].

With AIMS2, Meenan et al. [7] hoped to expand and strengthen the AIMS. To achieve the first goal, they added health status scales for arm function, support from family and friends, and work, and new satisfaction, attribution and prioritization scales for each of the 12 areas of health with which the new instrument deals. To achieve the second goal, they deleted or replaced the weak items of the original questionnaire and provided all items, including the visual analogue scale (VAS), with five-element response sets that could be transformed uniformly into discrete scales, rephrasing where necessary. The new instrument was shown to be more comprehensive and promises to be more sensitive to change than its predecessor, without loss of reliability or validity.

DUTCH-AIMS2 is a Dutch version of AIMS2. We developed it by making analogous adaptations to its predecessor DUTCH-AIMS and translating all new items from its original. Here, we report on its reliability and validity.

RESPONDENTS AND METHODS

Respondents

Ten rheumatologists from three hospitals asked out-patients who had definite or classical RA, and were not excluded by the following criteria, to participate in a survey on volunteer aid for arthritis patients. The exclusion criteria were residence in a nursing home, disease duration of <5 yr and age <16 yr. In each hospital out-patient rheumatology clinic, consecutive male and female patients were asked until 40 women had agreed to participate, after which only consecutive male patients were asked until a total of 40 men had agreed. Each patient was asked to name someone of the same age and sex, and from the same neighbourhood, who did not have RA and might be willing to participate in the survey (controls). Each participant's sex, age and level of education were recorded, together with the nature of any health problems other than RA and, in the case of the RA patients, disease duration.

The hospitals are in widely separated parts of The Netherlands; one serves an urban population, one a largely rural population and one a mixed population. A composite questionnaire was posted to each participant. Two weeks later, the questionnaires were collected by an assistant, who helped to complete them if necessary.

Questionnaire

The full questionnaire comprised the DUTCH-AIMS2 questionnaire, the Dutch version of the Modified Health Assessment Questionnaire (MHAQ) [8, 9], visual analogue scales for pain (VAS pain), fatigue (VAS fatigue) and loneliness (VAS loneliness), the Social Support List–Interactions (SSL12-I) [10] and
a scale to measure the extent to which social support is problematic (SS-neg) [11].

The DUTCH-AIMS2 is a translation of the AIMS2. For a detailed description of the AIMS2, we refer to Meenan et al. [7]. We used the original translation of the DUTCH-AIMS if relevant because this translation has already proven to be satisfactory. The new sections and items were translated by one of the authors (RPR). Next, two rheumatologists and one psychologist compared the translation with the original text to verify the cross-cultural equivalence.

The VAS scales for pain, fatigue and loneliness ranged from 0 to 100, where 0 corresponded to 'no pain', 'not tired at all' and 'not lonely at all', and 100 to 'severe pain', 'very tired', 'utterly lonely'.

SSL12-I comprises 12 questions on daily support, support with problems and esteem support. The possible responses are ranged on a scale of 1 (seldom or never) to 4 (often) and their average is the SSL 12-1 score [10]. SS-neg comprises four questions regarding unsupportive behaviour from friends or relatives, e.g. 'How often do people become annoyed when you don’t accept their advice?' Again, the score is the average of the scores for the individual items, which range from 1 (seldom or never) to 4 (often) [11].

The functional classification [12] and ACR 1987 criteria [13] of all participating patients from one hospital were assessed by a rheumatologist (JJR). In addition, the following laboratory parameters were assessed for these patients: ESR (Westergren method), haemoglobin (Hb), RF (Waaler–Rose test, positive when $>1:32$, or nephelometric assay, IgM-RF, positive when $>8$ IU). Laboratory results were used only if they were obtained no more than 14 days before or after completion of the questionnaire by the patient concerned.

Statistics

We assessed the reliability of the composite DUTCH-AIMS2 scales for the patient and control subpopulations in terms of internal consistency, which we measured with Cronbach’s $\alpha$ [14]. Values of $\alpha \geq 0.60$ indicate sufficient reliability for research purposes [15]. Following Meenan et al. [7], we chose a threshold value of 0.70. We also calculated $\alpha$ for MHAQ, SSL12-I and SS-neg.

Within-scale principal component factor analysis was performed for both groups. These analyses were repeated for the seven subpopulations formed by partitioning the patient group according to sex, according to age (< and $\geq 65$ yr) and according to level of education (primary school, junior vocational school and senior vocational school or higher).

We assessed the validity of the health status scales by several means.

1. The differences between the mean scores of the patient group and those of the control group were measured by analysis of covariance, controlled for age, sex and level of education.

2. Principal component factor analysis with varimax rotation was applied to the scores of the patients.

3. In two other evaluations, standards provided by the DUTCH-AIMS2 itself were used. For both, patients were divided into two groups. The differences between the two groups were tested with Student's $t$-test. For the first set of comparisons, the groups were defined by whether or not the corresponding area of health was a problem, as indicated in the attribution section. For the second, they were defined by whether or not the corresponding area had been chosen in the prioritization section.

4. Validity was further assessed on the basis of data independent of DUTCH-AIMS2. We calculated the correlation coefficients for the association between the patient scores on each of the health status scales and functional class, number of ACR 1987 criteria met, the laboratory parameters, MHAQ score, SSL12-I score, SS-neg score and the scores on each of the three VAS.

5. The validity of the individual satisfaction scales was assessed by comparing the scores of patients who had indicated in the attribution section that they have problems in the corresponding areas with those of patients who had not. For each area of health, we calculated the coefficients of correlation between the patients’ scores on the corresponding satisfaction scale and their scores on the corresponding health status scale, and also between their scores on the composite satisfaction scale and those on the arthritis impact scale.

To derive a clearer picture of the effects of arthritis on health status in our population, adjusted health status scales were derived by adjusting scores in accordance with the results of the attribution section. A health status score was multiplied by 0.50 if the patient had indicated that there were problems in the area concerned, and that these were due partially to arthritis and partially to other causes. Similarly, a score was multiplied by 0.25 if the problems were indicated as being entirely or mostly due to other causes. Otherwise, the scores were left unchanged.

RESULTS

Two hundred and thirty-one patients returned correctly completed questionnaires. Women made up 62% of the total patient population, but two hospitals did succeed in recruiting equal numbers of men and women. The mean age ($\pm$ s.d.) of the patients was 63.5 ($\pm$ 11.7) yr. Mean disease duration ($\pm$ s.d.) was 19.0 ($\pm$ 11.0) yr. The highest level of education completed was primary school for 38% and junior vocational school for 52%. Ten per cent had been educated to senior vocational school level or higher. Fifty-two per cent of the patient group had no co-morbidities. Seventeen per cent had hypertension, 15% had heart disease, 10% had lung disease and 11% had stomach problems.

One hundred and thirty-one controls, 68% of whom
were women, returned correctly completed questionnaires. The mean age (±S.D.) of this group of respondents was 59.5 (±12.8) yr. The highest level of education completed was primary school for 23% and junior vocational school for 48%. Twenty-nine per cent had been educated to senior vocational school level or higher. Morbidities were absent in 60%. Eighteen per cent suffered from hypertension, 15% from lung disease and 32% from other health problems.

Mean health status scores ranged from 0.15 on the self-care scale for control subjects to 6.53 on the walking and bending scale for subjects with RA. For the RA group, mean scores on eight health status scales were <5.0, but the ranges of the responses and corresponding scores were broad. The ranges and means of the DUTCH-AIMS2 health status scores of patients and controls are presented in Table I.

The average scores of the patient group on the satisfaction scales ranged from 1.9 (social support) to 3.8 (work).

Reliability
With one exception, the values of a for all health status scales exceeded 0.70 for both patients and controls (Table I). Most were in the range of 0.80–0.89. The exception is the internal consistency of the social activity scale for patients. Alpha was also below the threshold value for the social activity scale for five of the demographically defined subpopulations of patients, for the walking and bending scale for patients with no more than primary school education (a = 0.69), and for the work scale for two subpopulations with <23 members. The remaining values were similar to those obtained for corresponding scales for the patient group as a whole and for controls. In each of these cases and in the case of the level of tension scale for the control group, the percentage of the variance explained by the principal factor was ≥ 50%. Loading was also found to be on two factors in the social activities scale for six subpopulations of the patient group and in the work scale for patients aged ≥65 yr, of whom there were 23. Loading was on a single factor in all other cases, which together with those of the patient and control groups made up 98 (92%) of the 107 cases in which the analysis could be performed.

Validity
There was no significant difference between the mean scores of the patients and those of the controls on the health status scale of support from family and friends (F = 0.83; P > 0.05). On the other scales, the mean scores of the patients were significantly poorer (P < 0.001).

Principal component factor analysis was performed on 11 health status scales. The scores on the work scale were omitted from the analysis, because the number of respondents was insufficient. For the remaining scores, the analysis identified three factors, one mirroring physical dimensions: mobility (factor loading 0.80), walking and bending (0.79), hand and finger function (0.82), arm function (0.86), self-care (0.85) and household tasks (0.85); one mirroring psychological dimensions: level of tension (0.91) and mood (0.84);

<table>
<thead>
<tr>
<th>DUTCH-AIMS2 scale</th>
<th>Actual score range*</th>
<th>Mean ± s.d.</th>
<th>Internal consistency†</th>
<th>Actual score range*</th>
<th>Mean ± s.d.</th>
<th>Internal consistency†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>0-9</td>
<td>3.35 ± 2.33</td>
<td>0.80</td>
<td>0-6</td>
<td>0.73 ± 1.20</td>
<td>0.72</td>
</tr>
<tr>
<td>Walking and bending</td>
<td>0-10</td>
<td>6.53 ± 2.28</td>
<td>0.76</td>
<td>0-9</td>
<td>1.99 ± 2.29</td>
<td>0.75</td>
</tr>
<tr>
<td>Hand and finger function</td>
<td>0-10</td>
<td>4.52 ± 2.58</td>
<td>0.86</td>
<td>0-7</td>
<td>0.38 ± 1.01</td>
<td>0.79</td>
</tr>
<tr>
<td>Arm function</td>
<td>0-10</td>
<td>3.60 ± 2.62</td>
<td>0.84</td>
<td>0-6.5</td>
<td>0.21 ± 0.81</td>
<td>0.75</td>
</tr>
<tr>
<td>Self-care</td>
<td>0-10</td>
<td>2.23 ± 2.52</td>
<td>0.85</td>
<td>0-6.88</td>
<td>0.15 ± 0.88</td>
<td>0.83</td>
</tr>
<tr>
<td>Household tasks</td>
<td>0-10</td>
<td>4.20 ± 3.34</td>
<td>0.87</td>
<td>0-10</td>
<td>0.42 ± 1.46</td>
<td>0.87</td>
</tr>
<tr>
<td>Social activities</td>
<td>2.5-9.5</td>
<td>5.63 ± 1.22</td>
<td>0.66</td>
<td>0-7</td>
<td>4.69 ± 1.29</td>
<td>0.72</td>
</tr>
<tr>
<td>Support from family and friends</td>
<td>0-10</td>
<td>3.29 ± 2.64</td>
<td>0.88</td>
<td>0-10</td>
<td>2.72 ± 2.90</td>
<td>0.93</td>
</tr>
<tr>
<td>Arthritis pain</td>
<td>0-10</td>
<td>6.06 ± 2.06</td>
<td>0.82</td>
<td>0-7.5</td>
<td>1.56 ± 2.18</td>
<td>0.91</td>
</tr>
<tr>
<td>Work†</td>
<td>0-8.75</td>
<td>5.18 ± 2.41</td>
<td>0.76</td>
<td>0-8.75</td>
<td>1.33 ± 1.94</td>
<td>0.71</td>
</tr>
<tr>
<td>Level of tension</td>
<td>0-9.5</td>
<td>4.10 ± 1.99</td>
<td>0.89</td>
<td>0-8</td>
<td>2.92 ± 1.76</td>
<td>0.82</td>
</tr>
<tr>
<td>Mood</td>
<td>0.5-7.5</td>
<td>3.11 ± 1.49</td>
<td>0.78</td>
<td>0-7.5</td>
<td>1.96 ± 1.46</td>
<td>0.82</td>
</tr>
</tbody>
</table>

*Potential range of scores on all scales is 0-10.
†Internal consistency estimated by Cronbach's coefficient a.
‡n = 76 patients and 74 for controls.
and one mirroring social dimensions: social activities (0.65) and support from family and friends (0.88). Pain had high loadings on two factors: physical (0.65) and psychological (0.49).

For each health status scale, the mean score of patients who indicated that the corresponding area of health was a problem was significantly poorer \((P \leq 0.001)\) than that of patients who did not. The mean scores for patients indicating the scale as a problem area were between 7.03 for walking and bending and 3.07 for self-care; mean scores for the other patients were between 5.05 for social activities and 0.66 for self-care.

On seven scales, patients who designated these scales as priority areas for health status improvement had significantly poorer scores compared to patients who did not. These seven scales are: mobility, walking and bending, arm function, self-care, support from family and friends, arthritis pain, and level of tension.

Choosing a particular area of health as a priority area for improvement was not associated with significantly worse health status on the corresponding health status scale in all cases. Patients who did not include household tasks or social activities among their priorities for improvement had slightly worse health status scores in the area concerned than those who had; however, these differences were not significant. Most patients who included household activities among their priorities for improvement were women (34 women and five men).

The areas most frequently included among the priorities for improvement were: arthritis pain, selected by 74\% \((n = 166)\) of patients; walking and bending, selected by 52\% \((n = 113)\); hand and finger function, selected by 41\% \((n = 93)\); mobility, selected by 37\% \((n = 82)\). The other areas were each selected by < 20\% of respondents.

Mean patient scores on the individual satisfaction scales ranged from 2.3 to 4.0 for those who indicated having problems in the corresponding areas and from 1.4 to 2.5 for those who did not. The differences were statistically significant in all cases \((P \leq 0.001\) for all 12). The correlation coefficients for patients' scores on the satisfaction scales with respect to their scores on the health status scales ranged from 0.45 to 0.70. The correlation coefficient between their scores on the composite satisfaction scale and those on the arthritis impact scale was 0.55.

In the attribution section, the number of subjects who indicated that they had a problem in the given areas of health ranged from 101 (support from family and friends) to 215 (arthritis pain).

Adjusting patient scores on the basis of the attribution scales has only a minor effect on health status scores. The greatest changes were in the level of tension and social activities scales, and smallest in the self-care and arm function scales.

The correlations between the six physical health status scales, the overall physical scale, and the scale for arthritis pain on the one hand and functional classification and ACR classification on the other, were all significant (Table II). With the exception of the scale for hand and finger function, the same DUTCH-AIMS2 scales were also significantly correlated with ESR (Table II). Only the mobility and arm function scales were correlated with Hb (Table II). RF was significantly correlated with the scales for hand and

### Table II

<table>
<thead>
<tr>
<th>DUTCH-AIMS2 scale</th>
<th>Functional class</th>
<th>ACR class</th>
<th>ESR</th>
<th>Hb</th>
<th>RF</th>
<th>VAS fatigue</th>
<th>VAS pain</th>
<th>MHAQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
<td>(n)</td>
</tr>
<tr>
<td>Mobility</td>
<td>0.38†</td>
<td>0.24§</td>
<td>0.33†</td>
<td>-0.25§</td>
<td>0.06</td>
<td>0.33†</td>
<td>0.36†</td>
<td>0.68†</td>
</tr>
<tr>
<td>Walking and bending</td>
<td>0.47†</td>
<td>0.37†</td>
<td>0.21§</td>
<td>-0.06</td>
<td>0.02</td>
<td>0.33†</td>
<td>0.45†</td>
<td>0.67†</td>
</tr>
<tr>
<td>Hand and finger function</td>
<td>0.47†</td>
<td>0.34†</td>
<td>0.16</td>
<td>-0.20</td>
<td>0.30†</td>
<td>0.36†</td>
<td>0.28†</td>
<td>0.67†</td>
</tr>
<tr>
<td>Arm function</td>
<td>0.47†</td>
<td>0.25§</td>
<td>0.34†</td>
<td>-0.26§</td>
<td>0.24§</td>
<td>0.37†</td>
<td>0.39†</td>
<td>0.75†</td>
</tr>
<tr>
<td>Self-care</td>
<td>0.47†</td>
<td>0.28§</td>
<td>0.31†</td>
<td>-0.18</td>
<td>0.22§</td>
<td>0.19†</td>
<td>0.24†</td>
<td>0.77†</td>
</tr>
<tr>
<td>Household tasks</td>
<td>0.48†</td>
<td>0.30§</td>
<td>0.22‡</td>
<td>-0.08</td>
<td>0.17</td>
<td>0.29†</td>
<td>0.29†</td>
<td>0.70†</td>
</tr>
<tr>
<td>Arthritis pain</td>
<td>0.31§</td>
<td>0.21</td>
<td>0.28§</td>
<td>-0.11</td>
<td>0.06</td>
<td>0.51†</td>
<td>0.66†</td>
<td>0.40†</td>
</tr>
<tr>
<td>Physical</td>
<td>0.56†</td>
<td>0.44†</td>
<td>0.25§</td>
<td>-0.16</td>
<td>0.23§</td>
<td>0.38†</td>
<td>0.39†</td>
<td>0.86†</td>
</tr>
</tbody>
</table>

*Pearson's correlation coefficients, except for functional classification and ACR classification, where the Spearman correlation coefficient is used.
† \(P \leq 0.001\).
§ \(P \leq 0.005\).
‡ \(P \leq 0.01\).
‰ \(P \leq 0.05\).
fingert function, arm function, and for self-care and with the composite physical scale. The correlations between these DUTCH-AIMS2 scales and MHAQ, VAS fatigue and VAS pain were likewise all significant (Table II).

The scales for social activities, support from family and friends, the overall social interactions scale, the scale for level of tension, mood and the composite psychological scale each correlated significantly with both VAS fatigue and VAS loneliness (Table III). Each of the social scales was significantly correlated with SSL12-I and with SS-neg; the psychological scales were significantly correlated with SS-neg only (Table III).

**DISCUSSION**

AIMS2 has some advantages over AIMS1. Three new scales are added to assess arm function, support from family and friends and work. Work problems were not assessed in the AIMS1. The arm function and support scales do not introduce new health status areas, but complement the hand and finger function and social activities scales. The addition of a satisfaction section represents an increase in comprehensiveness over AIMS1. It has been shown that satisfaction with health status can vary among patients with the same level of health status [16]. The inclusion of a section to measure the attribution of arthritis health status problems to arthritis or other causes is important because many patients, especially in subject groups with elderly patients, have co-morbid conditions that can independently affect their health status. It has been argued that outcome assessment should be focused on those aspects of health status that are of most concern to the patient [17]. Therefore Meenan et al. [7] introduced a prioritization section in the AIMS2. Patients are asked to designate their top three priority areas for improvement. Finally, all questions in the AIMS2 were provided with five response options to produce a standardized format for all items. Using five response options should also produce measures that are more sensitive to improvements than would be possible with the yes/no options that were used in some of the AIMS1 scales.

The results of this initial assessment of the DUTCH-AIMS2 show that the reliability of the revised and expanded version is generally stable across two major subpopulations (RA patients and non-sufferers) and three major demographic dimensions (sex, age and level of education). We only included patients with disease duration of >4 yr, so we cannot extend our conclusions to all RA patients. The sensitivity to change and the test–retest reliability of the DUTCH-AIMS2 were not part of this study.

The results presented here are not directly comparable to those of studies concerning DUTCH-AIMS, owing to the differences in the study populations that resulted from the inclusion criteria. It is, therefore, uncertain to what extent corresponding scales would elicit similar scores. The new health status scales are, however, as reliable as those that had been carried over in modified form from DUTCH-AIMS.

For translation, we applied the items of the DUTCH-AIMS, which have proven to be valid [5]. New items could easily be translated without cultural differences.

Meenan and Mason [18] identified three dimensions for the variables represented by the AIMS2 health status scales, with arthritis pain and work as separate dimensions. These were a physical dimension consisting of the health status scales for mobility, walking and bending, hand and finger function, arm function, self-care and household tasks; a psychological dimension consisting of the scales for level of tension and mood; and a social dimension consisting of the scales for social activities and support from family and friends. We identified similar dimensions for DUTCH-AIMS2 using principal component factor analysis. However, pain was not shown to be a separate dimension, but loaded on two factors: one mirroring a physical dimension and one mirroring a psychological dimension. This is in accordance with the view that pain is a multi-dimensional concept [19].

Satisfaction is correlated with health status, but not identical to it. These new items measure a largely distinct aspect of health status that can be assessed independently of the health status scales. Meenan et al. [7] drew similar conclusions and recommended that the composite scale be used as this distinct aspect of outcome in arthritis intervention studies.

For most areas of health status, health status scores were worse for patients selecting it as a priority for improvement than the scores of those who did not. However, social activities was one of the areas for which the health status scores of those selecting it as a priority for improvement were better than the scores of those who did not, and household activities was the other. Of all the health status scales, that for social activities was the least reliable, and the majority of those who included household activities among their priorities were women. The latter form a subpopulation with fewer functional limitations (MHAQ) than most, generally lower scores for loneliness and for the most part better scores on all other DUTCH-AIMS2 scales.

The DUTCH-AIMS2 has satisfactory validity, as shown by the results of the attribution and prioritization sections, the comparisons with the results of the controls and the correlations with external measures.

The physical health status scales are valid, as shown by their strong correlations with the majority of the external measures of health status, particularly the MHAQ functional disability scale. Among the weaker correlations in this regard were those with ESR and Hb. The same had been found for the corresponding scales of DUTCH-AIMS [5]. The most likely explanation is that ESR, and to a lesser extent Hb, is a measure of disease activity rather than health status as such, whereas the physical health status scales address disease consequences. This subsection of DUTCH-AIMS2 is, therefore, a suitable measure of functional disability.
The validity of the social health status scales was indicated by moderate to strong correlations with the external measures of social support, and that of the psychological health status scales by their strong correlations with the external measures of fatigue and loneliness.

DUTCH-AIMS2 is more comprehensive than its predecessor, and is acceptably reliable and valid, both internally and externally, for use in clinical trials and outcome research in a variety of populations.

ACKNOWLEDGEMENTS

This study was supported by grants from the Nationaal Reumafonds, and the Ministry of Health, Welfare and Sport of the Netherlands. We wish to thank the following rheumatologists: H. J. Bernelot Moens, G. A. W. Bruyn, J. J. M. Festen, E. N. Griep, M. W. M. Kruisjen, M. A. F. J. van de Laar, J. C. M. Oostveen and J. M. G. W. Wouters for their help in selecting the patients.

REFERENCES


