**EXTENDED REPORT**

Assessment of disability with the World Health Organisation Disability Assessment Schedule II in patients with ankylosing spondylitis

A van Tubergen, R Landewé, L Heuft-Dorenbosch, A Spoorenberg, D van der Heijde, H van der Tempel, S van der Linden

**Objective:** To investigate in ankylosing spondylitis (AS) whether the newly developed World Health Organisation Disability Assessment Schedule II (WHODAS II) is a useful instrument for measuring disability, to assess its responsiveness in relation to other traditional disease specific instruments, and to identify factors that are associated with both short term and long term scores on the WHODAS II.

**Methods:** Patients with AS from a randomised controlled trial assessing the efficacy of spa treatment (n=117) and from a five year longitudinal observational study (n=97) participated. The patients completed several questionnaires, including the WHODAS II. After a three week course of spa treatment, 31 patients again completed all questionnaires to assess responsiveness. To determine to what degree the WHODAS II reflects some AS oriented measures on disease activity, functioning, and quality of life, correlation coefficients between the WHODAS II and these other questionnaires were calculated. Responsiveness was calculated by the effect size (ES) and standardised response mean (SRM). Linear regression analysis was performed to explore which factors might be associated with short term changes on the WHODAS II and to investigate (in the observational study) which factors of WHODAS II might predict disability five years later.

**Results:** Mean score on the WHODAS II was 23.9 (SD 15.5 (range 0.0–76.1)). Scores on the WHODAS II were significantly correlated with all disease specific questionnaires measured (all p<0.001). The WHODAS II showed a comparable short term responsiveness score (SRM 0.41; ES 0.39). In regression analysis these short term changes on the WHODAS II were significantly associated with changes in functioning (β coefficient 4.25, 95% confidence interval (95% CI) 1.24 to 7.26, p=0.007). In the observational study, disease activity (β coefficient 0.35, 95% CI 0.17 to 0.53, p<0.000) as well as functioning (β coefficient 0.23, 95% CI 0.09 to 0.38, p=0.002) seemed to significantly predict disability (WHODAS II) after five years.

**Conclusion:** The WHODAS II is a useful instrument for measuring disability in AS in that it accurately reflects disease specific instruments and that it shows similar responsiveness scores. In AS, a short term change on the WHODAS II is associated with a change in physical function. At the group level, disease activity and physical functioning may predict disability after five years.

**Abbreviations:** AP, anteroposterior; AS, ankylosing spondylitis; ASQoL, AS Quality of Life Questionnaire; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASRI, Bath Ankylosing Spondylitis Radiology Index total; CRP, C reactive protein; DFI, Dougados Functional Index; EQ-5D, EuroQol; ES, effect size; ESR, erythrocyte sedimentation rate; HAGS, Health Assessment Questionnaire for Spondyloarthropathies; ICF, International Classification of Functioning, Disability, and Health; ICIDH, International Classification of Impairments, Disabilities, and Handicaps; OASIS, study on outcome in AS; SF-36, Short Form-36 Health Status Questionnaire; SRM, standardised response mean; VAS, visual analogue scale; WHODAS II, World Health Organisation Disability Assessment Schedule II.
Assessment of disability with the WHODAS II in patients with ankylosing spondylitis

as oriented instruments that may predict outcome on study was to investigate at the group level whether there are
dependent on structural changes. The last aim of our
tors and factors that are subject to short term fluctuations, but
WHODAS II would be independent of sociodemographic fac-
tervention. We hypothesised that a short term change on the
second validation step, we assessed which, if any, disease spe-
(71%) Dutch patients participating in an ongoing international longitudinal observational study on outcome in AS (OASIS).
They completed the WHODAS II, some AS oriented questionnaires, and the SF-36, and underwent physical examination, laboratory tests, and radiographs of the axial skeleton in the summer of 2001. Patients from this study have been followed up since the summer of 1996. During this five year follow up period 40 patients withdrew from the observational study for various reasons.

**Methods**

All patients completed reports themselves containing socio-
demographic questions and a number of (AS oriented) ques-
tionnaires (see below). We used the self administered 36 item version of the WHODAS II in Dutch. The WHODAS II consists of 36 Likert formatted questions, divided into six domains. The final score is calculated with an SPSS syntax (available through the WHO). We used two different syntax versions—the 36 version if the patient had a job and the 32 version if patients did not have a job. In the second version, the four questions on work ability are omitted. The final scores range from 0 (best) to 100 (worst).

Physical function was measured with the Bath Ankylosing Spondylitis Functional Index (BASFI), the Dougados Functional Index (DFI), and the Health Assessment Questionnaire for Spondyloarthropathies (HAQ-S). The BASFI contains 10 questions on activities of daily living, and is scored on a 10 cm visual analogue scale (VAS), with anchors “easy” and “impossible” at either side. The mean of the items defines the final score, with scores ranging from 0 (best) to 10 (worst). The DFI consists of 20 Likert formatted scales, and includes activities of daily living. Originally, the DFI was completed on a three point Likert scale, and later it was modified to a five point Likert scale. In the present study we applied the five point Likert scale. For each item possible scores are 0, 0.5, 1, 1.5, and 2. Total scores range from 0 (best) to 40 (worst). The HAQ-S consists of eight subscales on health status and physical function, and is extended with five additional AS specific items on physical function, divided into two subscales. The questions are answered on Likert formatted scales. The mean of all subscales defines the final score, ranging from 0 (best) to 3 (worst). The HAQ-S was given to the patients from the spa treatment trial only.

**PATIENTS AND METHODS**

**Patients**

Altogether 214 patients with AS from two different studies participated. Study group A consisted of 117 out of 120 (98%) Dutch outpatients who had participated in a randomised controlled trial to assess the efficacy of a three week course of spa exercise treatment in AS. They completed the WHODAS II, some AS oriented questionnaires, and the SF-36 at the end of the study period (January 2000). From this trial it was concluded that spa exercise treatment was efficacious for patients with AS.

Patients from the control group of this trial were also offered a three week course of spa treatment after completion of the study period. Thirty one patients from this group completed all questionnaires a second time one week after the end of the spa treatment to assess responsiveness of these questionnaires. Study group B consisted of 97 out of 137 (71%) Dutch patients participating in an ongoing international longitudinal observational study on outcome in AS (OASIS). They completed the WHODAS II, some AS oriented questionnaires, and the SF-36, and underwent physical examination, laboratory tests, and radiographs of the axial skeleton in the summer of 2001. Patients from this study have been followed up since the summer of 1996. During this five year follow up period 40 patients withdrew from the observational study for various reasons.

**Methods**

All patients completed reports themselves containing socio-
demographic questions and a number of (AS oriented) ques-
tionnaires (see below). We used the self administered 36 item version of the WHODAS II in Dutch. The WHODAS II consists of 36 Likert formatted questions, divided into six domains. The final score is calculated with an SPSS syntax (available through the WHO). We used two different syntax versions—the 36 version if the patient had a job and the 32 version if patients did not have a job. In the second version, the four questions on work ability are omitted. The final scores range from 0 (best) to 100 (worst).

Physical function was measured with the Bath Ankylosing Spondylitis Functional Index (BASFI), the Dougados Functional Index (DFI), and the Health Assessment Questionnaire for Spondyloarthropathies (HAQ-S). The BASFI contains 10 questions on activities of daily living, and is scored on a 10 cm visual analogue scale (VAS), with anchors “easy” and “impossible” at either side. The mean of the items defines the final score, with scores ranging from 0 (best) to 10 (worst). The DFI consists of 20 Likert formatted scales, and includes activities of daily living. Originally, the DFI was completed on a three point Likert scale, and later it was modified to a five point Likert scale. In the present study we applied the five point Likert scale. For each item possible scores are 0, 0.5, 1, 1.5, and 2. Total scores range from 0 (best) to 40 (worst). The HAQ-S consists of eight subscales on health status and physical function, and is extended with five additional AS specific items on physical function, divided into two subscales. The questions are answered on Likert formatted scales. The mean of all subscales defines the final score, ranging from 0 (best) to 3 (worst). The HAQ-S was given to the patients from the spa treatment trial only.

**Assessment of disability with the WHODAS II in patients with ankylosing spondylitis**

Assessment of disability with the WHODAS II in patients with ankylosing spondylitis 141

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the study groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total population</td>
</tr>
<tr>
<td></td>
<td>(n=214)</td>
</tr>
<tr>
<td>Male [%]</td>
<td>69</td>
</tr>
<tr>
<td>Age*</td>
<td>50 (11, 23–82)</td>
</tr>
<tr>
<td>Time since diagnosis*</td>
<td>14 (8, 2–46)</td>
</tr>
<tr>
<td>Duration of complaints*</td>
<td>23 (11, 3–56)</td>
</tr>
<tr>
<td>Patients with a (un)paid job [%]</td>
<td>63</td>
</tr>
</tbody>
</table>

*Mean (SD, range) years. OASIS, outcome in ankylosing spondylitis international study.

**Table 2** Scores on the WHODAS II, AS oriented questionnaires, and the SF-36 in the total study population (n=214)

<table>
<thead>
<tr>
<th>Questionnaire (range)</th>
<th>Score (mean (SD))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHODAS II (0–100)</td>
<td>23.9 (15.5)</td>
</tr>
<tr>
<td>BASFI (0–10)</td>
<td>4.3 (2.3)</td>
</tr>
<tr>
<td>DFI (0–40)</td>
<td>10.9 (7.0)</td>
</tr>
<tr>
<td>HAQ-S (0–3)*</td>
<td>0.82 (0.55)</td>
</tr>
<tr>
<td>BASDAI (0–10)</td>
<td>3.8 (2.2)</td>
</tr>
<tr>
<td>Patient global (0–10)</td>
<td>4.2 (2.5)</td>
</tr>
<tr>
<td>Spinal pain (0–10)</td>
<td>4.2 (2.6)</td>
</tr>
<tr>
<td>ASQoL (0–18)</td>
<td>6.4 (4.8)</td>
</tr>
<tr>
<td>SF-36 (0–100)</td>
<td>64.6 (22.8)</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>73.1 (22.8)</td>
</tr>
<tr>
<td>Role limitation (physical)</td>
<td>50.1 (42.0)</td>
</tr>
<tr>
<td>Role limitation (emotional)</td>
<td>75.4 (37.6)</td>
</tr>
<tr>
<td>Mental health</td>
<td>72.9 (18.8)</td>
</tr>
<tr>
<td>Vitality</td>
<td>55.8 (21.3)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>58.1 (20.3)</td>
</tr>
<tr>
<td>General health</td>
<td>49.4 (19.9)</td>
</tr>
</tbody>
</table>

*Data from spa treatment trial only. AS, ankylosing spondylitis; WHODAS II, World Health Organisation Disease Ability Score II; BASFI, Bath Ankylosing Spondylitis Functional Index; DFI, Dougados Functional Index; HAQ-S, Health Assessment Questionnaire for Spondyloarthropathies; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; ASQoL, Ankylosing Spondylitis Quality of Life Questionnaire; SF-36, Short Form 36.
The scores were combined according to the Bath AS Radiology normal, suspicious, mild, moderate, and severe, respectively. The lumbar spine, cervical spine, and hips were graded 0 to 4 (for cervical spine) were taken. The radiographs of the sacroiliac joints, erythrocyte sedimentation rate (ESR) and C reactive protein (CRP) concentration. Radiographs of the axial skeleton were against the wall and the chin distance between the patient’s occiput and the wall when the lower sternum. Occiput to wall distance was measured as the difference in chest circumference between full inspiration and full expiration at the level of the lower sternum. Finger to floor distance was measured as the distance between the patient’s fingertips and the floor after forward flexion on a 10 cm segment marked on the patient’s back. The Schober test was consisted of the Schober index, finger to floor distance, chest expansion, and occiput to wall distance. The score of the WHODAS II, independent variables were change scores on the WHODAS II, physical function, emotional problem), mental health, vitality, bodily pain, and general health perceptions. Scores range from 0 to 100 (worst). Quality of life was measured with the AS quality of Life Questionnaire (ASQoL) and the SF-36. The ASQoL is a newly developed AS oriented quality of life instrument containing 18 yes or no questions. Scores range from 0 to 18 with lower scores implying a better quality of life. The SF-36 is a widely applied generic instrument for measuring health status and consists of eight domains: physical functioning, role functioning, role limitations (physical problem), role limitations (emotional problem), mental health, vitality, bodily pain, and general health perceptions. Scores range from 0 to 100 (worst).

The patients from the OASIS study were also given a physical examination, and laboratory and radiographic assessments at regular intervals. Measures for determining spinal mobility consisted of the Schober index, finger to floor distance, chest expansion, and occiput to wall distance. The Schober test was measured as the increase in the distance after maximal forward flexion on a 10 cm segment marked on the patient’s back. Finger to floor distance was measured as the distance between the patient’s fingertips and the floor after forward flexion with knees straight and feet together. Chest expansion was measured as the difference in chest circumference between full inspiration and full expiration at the level of the lower sternum. Occiput to wall distance was measured as the distance between the patient’s occiput and the wall when the patient’s heels and back were against the wall and the chin was held level. Laboratory variables assessed included erythrocyte sedimentation rate (ESR) and C reactive protein (CRP) concentration. Radiographs of the axial skeleton (pelvis, anteroposterior and lateral lumbar spine, lateral cervical spine) were taken. The radiographs of the sacroiliac joints, lumbar spine, cervical spine, and hips were graded 0 to 4 (for normal, suspicious, mild, moderate, and severe, respectively). The scores were combined according to the Bath AS Radiology Index total (BASRI-t) for the spine and hips. The score of the BASRI-t ranges from 0 (best) to 16 (worst). At physical examination, disease activity by both patient and doctor (on a 10 cm VAS) were also measured.

### Table 3 Correlations between WHODAS II, AS oriented questionnaires, and the domains of the SF-36

<table>
<thead>
<tr>
<th></th>
<th>SF-36</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHODAS II</td>
<td>Physical function</td>
</tr>
<tr>
<td>WHODAS II</td>
<td>0.70</td>
</tr>
<tr>
<td>BASFI</td>
<td>0.67</td>
</tr>
<tr>
<td>DFI</td>
<td>0.70</td>
</tr>
<tr>
<td>HAQ-S*</td>
<td>0.67</td>
</tr>
<tr>
<td>BASDAI</td>
<td>0.57</td>
</tr>
<tr>
<td>Pt Global</td>
<td>0.58</td>
</tr>
<tr>
<td>Spinal pain</td>
<td>0.54</td>
</tr>
<tr>
<td>ASQoL</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Correlation coefficients were calculated with Pearson’s correlation tests and are significant at the 0.01 level. *Data from spa therapy trial patients only.

### Figure 1

WHODAS II plotted against the BASFI and the ASQoL.

Data from the two different study groups were pooled and scores are presented using descriptive statistics. To assess whether the WHODAS II would reflect AS oriented questionnaires, Pearson’s correlations were used. The same exercise was done for the WHODAS II versus the SF-36, and as an external criterion, the SF-36 versus all AS oriented questionnaires. Scatter plots were made to visualise scores on the WHODAS II compared with scores on AS oriented questionnaires.

Responsiveness was assessed with data from the 31 patients who had received a three week course of spa treatment and completed a second questionnaire thereafter. Responsiveness was calculated by two different methods: the effect size (ES) and the standardised response mean (SRM). The ES was calculated as the mean change after treatment compared with baseline, divided by the standard deviation (SD) of the baseline scores. The SRM is similar to the ES, but with the SD of the change score as denominator. The results were interpreted according to the Cohen’s effect size index, in which 0.2 refers to a small change, 0.5 to moderate, and 0.8 or more to large change.

Pearson’s correlations between the WHODAS II and AS oriented questionnaires were also calculated on the change scores (difference in scores before and after the three week course of spa treatment). Forward stepwise selection linear regression analysis was performed to investigate which factors were associated with this short term change on the WHODAS II. The dependent variable was the change score on the WHODAS II, independent variables were change scores on the BASFI, BASDAI, spinal pain, patient global wellbeing, and
ASQoL. Adjustments were made for age, sex, and disease duration. Because the BASFI, DFI, and HAQ-S showed a high level of collinearity (tolerance <0.4), only the BASFI score was used as covariate, and the DFI and HAQ-S were excluded.

Forward stepwise linear regression analysis was also used to investigate which factors may predict disability, as measured by the WHODAS II, five years later. The 1996 baseline data from the OASIS study were used for this purpose. The dependent variable was the WHODAS II score. Independent variables were not only AS oriented questionnaires (BASDAI, BASFI, patient global assessment, and spinal pain), but also measures of physical examination (occupit to wall distance, finger to floor distance, Schober's test, chest expansion), laboratory values (ESR and CRP), scores on radiographs (BASRI), and disease activity by both patient and doctor. Adjustments were made for age, sex, and disease duration. Because the BASFI and DFI again showed a high level of collinearity (tolerance <0.4), only the BASFI score was used as a covariate, and the DFI was excluded. Occupit to wall distance, ESR, CRP, and disease activity by the doctor were non-normally distributed and therefore log transformed before introduction into the regression model.

RESULTS

Table 1 shows some characteristics of the study groups. Patients from the spa treatment trial showed lower disease duration (time since diagnosis) and duration of complaints compared with the OASIS group. Sixty three percent of the patients from the total study population had a paid or unpaid job. The scores of the WHODAS II in this group of patients from the total study population were therefore calculated with the 36 item version SPSS syntax, and the scores in the group of patients without a job were calculated with the 36 item version SPSS syntax.

Table 2 presents the final scores of the total study population on the WHODAS II and the results on the AS oriented questionnaires and the SF-36. The WHODAS II showed a mean score of 23.9 (SD 15.5; median 24.5; range 0.0 to 76.1). The other questionnaires showed comparable scores on the opposite direction.

The scores on the WHODAS II were significantly correlated with AS oriented questionnaires on physical functioning, disease activity, and quality of life (all p<0.001, table 3). The WHODAS II and all AS oriented questionnaires also correlated well with all domains of the generic health status questionnaire SF-36 (all p<0.001). Figure 1 shows two examples of scatter plots in which the WHODAS II is plotted against the BASFI and the ASQoL respectively.

Table 4 shows the scores on the questionnaires before and after spa treatment from the 31 patients who had a three week course of spa treatment, including the responsiveness scores. In all variables an improvement was seen, with moderate to large responsiveness scores. The WHODAS II showed a comparable, moderate responsiveness score (SRM 0.41 and ES 0.39).

Table 5 shows the correlation coefficients between the change scores on the WHODAS II and the change scores on the AS oriented questionnaires. In the short term, a change in WHODAS II was significantly correlated with a change in physical functioning (BASFI and DFI), but not with a change in disease activity or quality of life. With forward stepwise linear regression analysis the change score on the BASFI was selected as the only variable to be associated with short term change on the WHODAS II (adjusted R²=0.21, β coefficient for BASFI=4.25 (95% confidence interval (95% CI)1.12 to 7.26, p=0.007; β coefficient for intercept=1.59 (95% CI–3.19 to 6.37), p=0.501).

Forward stepwise linear regression analysis was also performed to investigate whether clinical and demographic variables could predict disability, as measured by the WHODAS II after a five year period. The variables BASDAI and BASFI at baseline were selected as independently associated with WHODAS II scores after five years (adjusted R²=0.48, table 6).
DISCUSSION

In this study a new generic instrument for measuring disability and functioning—the WHODAS II, which is conceptually compatible with the recently revised WHO ICF—has been investigated in patients with AS. The questionnaire was given to patients who had participated in an intervention trial and to patients participating in an ongoing longitudinal observational study. Both groups together showed a wide range of duration of disease and duration of complaints. The advantage of this approach is to enhance the ability to generalise the results. A large range in scores on the WHODAS II was found (0 to 76). A relevant ceiling effect of the WHODAS II score was not to be expected because—despite the problems in activities of daily living—patients with AS are seldom completely disabled, not even after a long duration of disease.18

Because it is a generic instrument, the WHODAS II would particularly be useful for studies designed to compare across different interventions and different disease entities (for example, in economic evaluations), similar to other generic instruments such as the SF-36 and the EuroQol (EQ-5D).19 20 However, the WHODAS II is different from the other measures of health status in that it has been based on an international classification system; it is applicable across cultures; and it can be applied as an instrument to measure overall disability at a disorder specific level, as we have investigated in AS. The bodily, personal, and social levels of disability are well represented in the WHODAS II in our study: scores on the WHODAS II were significantly correlated with all AS oriented questionnaires on disease activity, physical functioning, and quality of life. These findings were similar to the external criterion, the SF-36, thereby further validating the WHODAS II.

We also investigated which variables were associated with a short term change in the WHODAS II and which variables may predict disability over a five year period. A short term change in the WHODAS II was to some degree associated with a short term change in physical function (BASFI; R2=0.21), and not with a change in disease activity. This finding further strengthens the validation of the WHODAS II in that it is not influenced by measures that may show a daily fluctuating pattern (such as disease activity), but that it is influenced by factors that induce structural changes. Variables that best predict disability (WHODAS II scores) after five years were physical function (BASFI) and disease activity (BASDAI) (R2=0.48). Apparently, WHODAS II scores are reasonably associated with early predictors of outcome in AS, which further improves its quality as an outcome measure. However, it can be seen from the relatively low adjusted R2 that other factors which have not been assessed in our disease oriented study, may also influence disability (WHODAS II) in the long term.

A few methodological issues should be considered. We did not give the WHODAS II twice to a group of patients who remained stable. Therefore, we could not assess the reproducibility of the WHODAS II in AS. Future studies will be needed to establish this. However, in extensive testing, the WHODAS II has been shown to have a test-retest reliability ranging from 0.91 to 0.95 (within class correlations) across respondents from different geographical regions and from physical disorder, general population, drug, alcohol, and mental disorder samples.21 22 Secondly, as a consequence, we have no information on to what degree the responsiveness scores of the WHODAS II may be determined by a potential degree of lack of reproducibility. However, the level of responsiveness was similar to the AS oriented questionnaires that have been shown to be reproducible and responsive in previous studies.1–3 11 12–20 Therefore, we think that the present results reliably reflect clinical findings.

In patient populations with schizophrenia, depression, and alcohol dependence, the effect size of the WHODAS II was 1.46, 1.05, and 1.35 respectively, showing that it is able to detect even larger changes than the moderate, although comparable, effect size in our study.21 26

In conclusion, the WHODAS II is a useful instrument for measuring disability in AS in that it accurately incorporates results of AS oriented instruments on disease activity, physical functioning, and quality of life. It shows responsiveness scores that are comparable with AS oriented instruments. In AS, a short term change on the WHODAS II is partly associated with a change in physical function. Activity and physical functioning seem to be strong predictors of disability, as measured by the WHODAS II, after five years.

Authors’ affiliations
A van Tubergen, R Landewé, H Heuft-Dorenbosch, et al
Division of Rheumatology, University Hospital Maastricht, Maastricht, The Netherlands
H van der Tempel, Department of Rheumatology, Maasland Ziekenhuis Sittard, The Netherlands

REFERENCES


