Measurement of Spinal Mobility in Ankylosing Spondylitis: Comparison of Occiput-to-Wall and Tragus-to-Wall Distance

LIESBETH HEUFT-DORENBOSCH, DEBBY VOSSE, ROBERT LANDEWÉ, ANNEKE SPOORENBERG, MAXIME DOUGADOS, HERMAN MIELANTS, HILLE van der TEMPEL, SJEF van der LINDEN, and DÉSIRÉE van der HEIJDE

ABSTRACT. ***Objective.*** To investigate if the tragus-to-wall distance (TWD) is more reliable compared to the occiput-to-wall distance (OWD) as a measurement for thoracic spine extension in patients with ankylosing spondylitis (AS).

Methods. Data from the OASIS cohort, an international longitudinal observational study on outcome in AS, were used. Measurements of OWD and TWD were performed at baseline and at 6, 12, 18, and 24 months. Paired data of T_x and $T_{x+6months}$ were used to perform test-retest measurements (intraclass correlations, limits of agreement, and interperiod correlation matrix). Bland and Altman plots were constructed to investigate the agreement between both observations, assuming that there was no true change between 0 and 6 months. To investigate whether a change in disease activity would have influenced the results, limits of agreement were calculated in a subgroup of patients with a stable Bath Ankylosing Spondylitis Disease Activity Index (BASDAI; defined as a maximum BASDAI change of ± 1) between T_0 and T_6 and compared with the results of the whole group. Limits of agreement were also calculated for kyphosed patients only.

Results. The test-retest intraclass correlations were between 0.94 and 0.96 for OWD and between 0.93 and 0.95 for TWD. The direct measurement-remeasurement correlation calculated by extrapolation of the interperiod correlation regression line was 0.92 for OWD and 0.90 for TWD. OWD and TWD showed comparable reliability on the entire value of scores. The lower 95% limit of agreement was between –3.4 cm and –2.5 cm for OWD and between –3.4 cm and –3.1 cm for TWD. The upper limit of agreement was between 3.1 cm and 4.2 cm for OWD and between 2.9 cm and 3.9 cm for TWD. In all patients as well as in kyphosed patients only, limits of agreement were comparable between OWD and TWD. The patterns of the scatterplots according to Bland and Altman were similar for OWD and TWD. Measurement error was more pronounced in kyphosed patients compared to patients with a normal thoracic extension. However, over the entire range of kyphosis, measurement error was similar.

Conclusion. OWD and TWD are equally reliable in assessing thoracic spine extension. Although the TWD is in general easier to perform in AS patients compared to OWD, we recommend the OWD measurement over TWD: in OWD measurement a value of zero easily distinguishes patients with normal thoracic spine extension from kyphosed patients. (J Rheumatol 2004;31:1779–84)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS
SPINAL MOBILITY
OUTCOME MEASURES
THORACIC EXTENSION

Reduced spinal mobility and changes in posture are characteristic features of ankylosing spondylitis (AS). The typical patient with AS has a reduced lumbar lordosis and an increased thoracic kyphosis, and as a result the head is somewhat bent forward. In the course of the disease process these postural changes tend to progress and may become irreversible due to structural changes of the spine. To quantify thoracic kyphosis, the distance between occiput and wall (occiput-to-wall distance, OWD) is assessed when the patient is standing erect with stretched knees and the back

From the Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, Maastricht, The Netherlands.

Supported by a grant from the Dutch Arthritis Association.

L. Heuft-Dorenbosch, MD, Rheumatologist; D. Vosse, MD, Rheumatologist; R. Landewé, MD, PhD, Rheumatologist; S. van der Linden, MD, PhD, Professor in Rheumatology, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht; A. Spoorenberg, MD, Rheumatologist, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, Department of Rheumatology, Medical Center Leeuwarden, Leeuwarden, The Netherlands; D. van der Heijde, MD, PhD, Professor in Rheumatology, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, Department of Rheumatology, Limburg University Center, Diepenbeek, Belgium; M. Dougados, MD, Professor in Rheumatology, Department of Rheumatology, Hospital Cochin, Paris, France; H. Mielants, MD, PhD, Professor in Rheumatology, Department of Rheumatology, University Hospital Gent, Gent, Belgium; H. van der Tempel, MD, Rheumatologist, Department of Rheumatology, Maasland Hospital Sittard, The Netherlands.

Address reprint requests to Dr. L. Heuft-Dorenbosch, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands. E-mail: eheuf@sint.azm.nl

Submitted March 28, 2003; revision accepted April 7, 2004.

against the wall. OWD is a measure for thoracic spine extension[1] and was selected as a core set instrument by the ASsessment in Ankylosing Spondylitis (ASAS) working group. This is an international working group of clinical experts, clinical epidemiologists, representatives of the pharmaceutical industry, and representatives of patient associations. This core set was selected in 1998 and consists of different domains with specific instruments per domain, and is to be used in various settings in clinical studies in AS. To assess the domain of spinal mobility, the ASAS working group selected 3 instruments as core set instruments[2]. In addition to the OWD, these instruments are chest expansion and modified Schober test.

In clinimetrics, important features of an instrument are (construct) validity, reliability, and responsiveness[3]. Reliability (synonyms: reproducibility, repeatability) is an expression of the extent to which similar results are obtained on repeated applications of the same assessment technique, assuming no true interval change in the phenomenon under study[1].

Another way to assess thoracic spine extension is by measuring the tragus-to-wall distance (TWD). Although not included in the core set, this is part of the Bath Ankylosing Spondylitis Metrology Index (BASMI), which is also widely used[4,5].

There are some arguments that favor TWD over OWD measurement. Together with thoracic spine extension, involuntary flexion and extension of the head can occur. Flexion and extension of the head take place in the atlanto-occipital joint and the cervical spinal joints and can interfere with the measurement of both OWD and TWD. We hypothesized that compared to the measurement of TWD, the measurement of OWD would be more influenced by concomitant flexion or extension of the head. The explanation for this lies in the fact that, compared to the occiput, the tragus lies closer to the sagittal axis of the flexion or extension movements (Figure 1). On the other hand, TWD might be influenced by unintended rotation of the head. Second, the OWD is more difficult to measure since the patient's hair frequently obscures the view of the occiput. Third, with the patient standing erect against the wall, if the OWD is small, the observer's view of the ruler is difficult and prone to parallax problems. Because of these arguments we wanted to test the hypothesis that the TWD instrument would be more reliable than the OWD as a measurement for thoracic spine extension.

MATERIALS AND METHODS

Patients. In our study we used the data from the OASIS cohort, an international longitudinal observational study on outcome in AS with followup visits according to a fixed protocol. Data from this cohort have been reported[6]. Participating centers are the University Hospital, Maastricht, The Netherlands; the Maasland Hospital, Sittard, The Netherlands; Hospital Cochin, Paris, France; and University Hospital, Gent, Belgium; which are all secondary and tertiary referral centers. Consecutive outpatients with an established diagnosis of AS according to the modified New York criteria were enrolled in 1996 and followed thereafter.

For this study we used data from visits at baseline (T_0) and at 6 (T_6), 12 (T_{12}), 18 (T_{18}), and 24 months (T_{24}). On each study visit, all patients completed a number of questionnaires and underwent a clinical examination. The same investigator performed all clinical examinations per country. In addition to the core set variables propagated by the ASAS working group, several additional measurements were done, including the TWD.

Occiput-to-wall distance. The patient stands with heels and buttocks touching the wall behind and with the knees straight. The patient is asked how far back he/she can get the head, still keeping the chin in the normal position. In the straight position, the distance between the posterior convexity of the occiput and the wall is measured to the nearest 0.1 centimeter using a rigid ruler[1]. The better of 2 attempts is recorded.

Tragus-to-wall distance. The patient is positioned as in measurement of the OWD. The distance between the tragus and the wall is measured to the nearest 0.1 centimeter using a rigid ruler[1]. The better of 2 attempts is recorded.

In our study, measurement of TWD immediately followed measurement of OWD without repositioning the patient. However, we are comparing assessments with an interval of 6 months, and obviously there was repositioning between the first and second assessment of the OWD and the TWD. Therefore, measurement variability due to repositioning (including flexion and rotation of the head) is included in the overall measurement error.

Statistical analysis. We first assessed the mean, range, and standard deviation of both OWD and TWD. Because of the character of the disease, with slow disease progression, we assumed on a group level that no relevant changes in spinal mobility occurred during a time interval of 6 months, and thus changes in this time interval would be due to measurement error rather than to real changes in spinal mobility. As we did not perform a true test-retest measurement to assess the measurement error of both the OWD and TWD, we used the paired data of T_x and $T_{x+6months}$. To determine if it was appropriate to do so, we used several methods.

First, random-effect single-measure intraclass correlations (ICC; type 2.1) were calculated for T_x and $T_{x+6months}$. Second, an interperiod correlation matrix was constructed. By this method the intercorrelations of the measurement periods are plotted against the intervening time intervals. These correlations can be linearly related to time, which may be represented by a well-fitting regression line. Extrapolation of this line through the y-axis gives the direct measurement-remeasurement correlation, which may be interpreted as a quality measure of the measurement[7]. This was done twice, first for all patients, then only for patients not able to reach normal thoracic extension. Third, limits of agreement as defined by Bland and Altman (explained below) were calculated in all patients with a stable Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score (defined as a maximum BASDAI change of ± 1) between T_0 and T_6. The results of this analysis were compared with the results for the whole group.

To visualize the agreement between 2 observations and to check if observed differences are similar along the total range of scores, we plotted for all patients the difference between the score at T_0 and at T_6 of the OWD and the TWD against each patient's mean of the 2 scores. This graphic representation by means of a scatterplot of 2 measurements is called the Bland-Altman plot. In this method the difference between 2 observations (on the y-axis) is plotted against the mean of the same 2 observations (on the x-axis). Discrepancies between the observations are thus visualized as well as any possible relationship between the various parts of the scale and the corresponding measurement error. The advantage compared to a simple plot of the results of one method against the other is that in this latter method the data points will usually be clustered, and between-method differences are difficult to assess. Thereafter we calculated the limits of agreement for OWD and TWD based on the 95% limits of agreement method by Bland and Altman using the formula:

$$\overline{d} \pm 1.96 * sd_{diff}$$

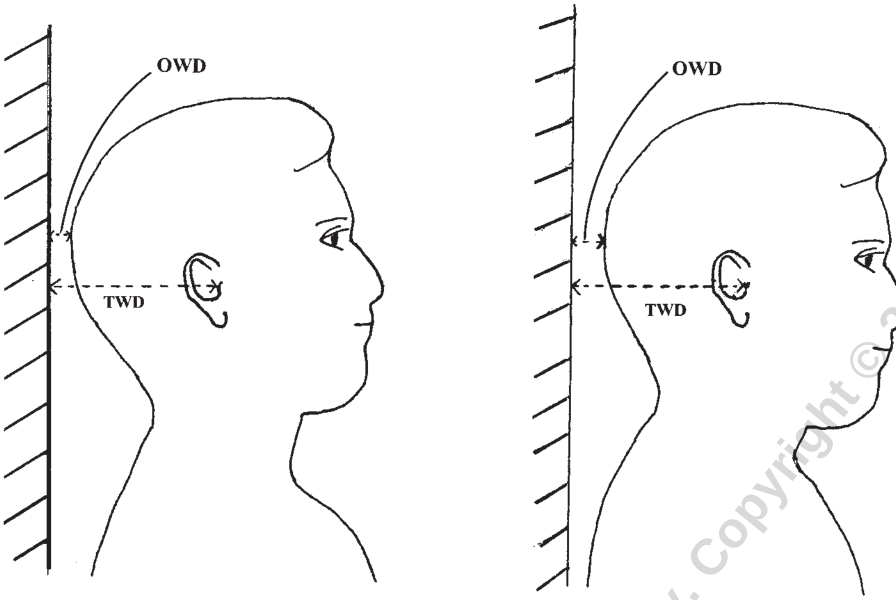

Figure 1. TWD in both positions of the head is the same, whereas the OWD is not, due to flexion of the neck.

where $\bar{d}$ represents the mean difference between the 2 observations and sd_{diff} the standard deviation of the difference[8]. Similar plots were made for the measurements comparing T_6 and T_{12}, T_{12} and T_{18}, T_{18} and T_{24}. In contrast with the OWD, which is zero in all patients with normal thoracic spine extension, the minimum TWD depends on the size of the head and the position of the ear. In patients able to touch the wall with their occiput, the possibility for measurement error due to patient variance seems to be smaller compared to patients not able to touch the wall, as the wall gives stability to the head. Thus a high proportion of patients with an OWD = 0 may result in smaller limits of agreement. In daily practice the assessment of both OWD and TWD will largely be used in kyphosed patients. Thus we also calculated the limits of agreement for kyphosed patients exclusively. To be able to compare TWD with OWD in the same patients, we selected all patients based on an OWD > 0 for calculations of limits of agreement for both OWD and TWD.

To assess whether the OWD and TWD instruments behave differently in different ranges of the scale, limits of agreement were calculated for the entire range of values, as well as for every quartile.

RESULTS

At baseline, 217 patients were included. Characteristics of the patients are presented as mean with standard deviation (SD), or as median with interquartile range if appropriate (Table 1). The test-retest ICC was between 0.94 and 0.96 for OWD and between 0.93 and 0.95 for TWD. The direct measurement-remeasurement correlation calculated by extrapolation of the regression line was 0.92 for OWD and 0.90 for TWD. With calculations only for kyphosed patients, these figures were 0.88 for OWD and 0.92 for TWD. Figure 2 shows this regression line for all patients.

Scatterplots according to Bland and Altman are shown for OWD and TWD for all patients for the comparison of T_0 and T_6 in Figures 3 and 4, respectively. The patterns of the plots were roughly similar for OWD and TWD. There is no influence of the magnitude of the OWD or TWD on the measurement error, i.e., the measurement error is similar along the entire scale. The summary results for the average scores and differences as well as the 95% limits of agreement comparing baseline and T_6, T_6 and T_{12}, T_{12} and T_{18}, and

Table 1. Baseline characteristics and scores on ASAS core set measures. Data are mean (SD) unless otherwise indicated.

	Total study population (n = 217)
Male/female	150/67
Age, yrs	43.1 (12.7)
Duration of complaints, yrs	19.6 (11.8)
Time since diagnosis, yrs	10.8 (8.9)
HLA-B27, present/absent/no data	155/32/30
History of IBD, present/absent/no data	17/145/55
History of uveitis, present/absent/no data	81/133/3
History of psoriasis, present/absent/no data	10/152/55
BASFI, score 0–10	3.4 (2.6)
VAS pain of the spine, 0–10 cm	3.5 (2.4)
Night pain, 4 point Likert Scale (IQR)	1.0 (1.0–2.0)
Chest expansion, cm	4.7 (2.2)
10 cm Schober, cm	2.8 (1.4)
Tragus-to-wall distance, cm*, (IQR)[range]	12.5 (11.0–16.0)[8.2–34.4]
Occiput-to-wall distance, cm*, (IQR)[range]	1.6 (0.0–6.0)[0.0–26.1]
VAS patient global, 0–10 cm	3.5 (2.8)
Peripheral arthritis, present/absent†	57/160
ESR, mm/h* (IQR)	10 (5–19)
CRP, mg/l* (IQR)	7 (6–19)
VAS physician on disease activity*, 0–10 (IQR)	1.4 (0.5–3.3)
VAS patient on disease activity, 0–10 cm	3.8 (2.8)
Duration of morning stiffness, min	36 (30)

* Median (interquartile range); IBD: inflammatory bowel disease; BASFI: Bath Ankylosing Spondylitis Functional Index; VAS: visual analog scale; IQR: interquartile range. † Defined as ≥ 1 swollen joint on physical examination.

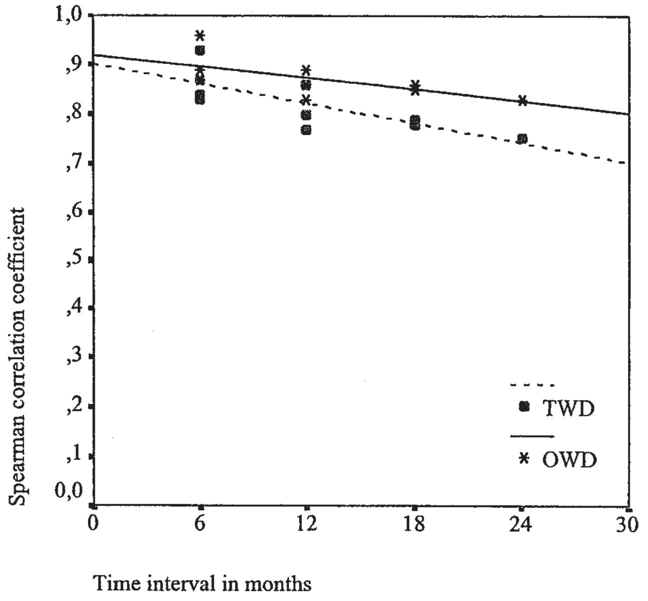

Figure 2. Graph of Spearman correlation coefficients against each interperiod with regression lines for OWD and TWD.

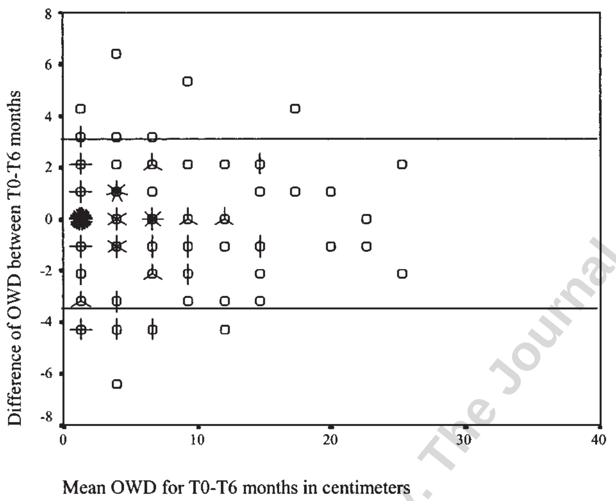

Figure 3. Graph of the difference scores against the mean score for OWD in cm in all patients. Upper line indicates the upper limit (+2 SD) and the lower line (–2 SD) the lower limit of the 95% of agreement interval. Each square represents one patient; each additional spike of the "sunflower" represents an additional patient (for example, the data point in the lower part of the figure below the line of –2 SD represents 5 patients).

T_{18} and T_{24} are shown in Table 2. Results were comparable between all paired observations. Comparing OWD and TWD at various time intervals, there is no consistent difference between the means and the SD in either the upper or the lower limit of agreement. As well, the limits of agreement for both OWD and TWD in all patients did not differ importantly. The lower 95% limit of agreement was between –3.4 cm and –2.5 cm for OWD and between –3.4 cm and –3.1 cm for TWD. The upper limit of agreement was between 3.1 cm and 4.2 cm for OWD and between 2.9 cm and 3.9 cm for TWD (Table 3).

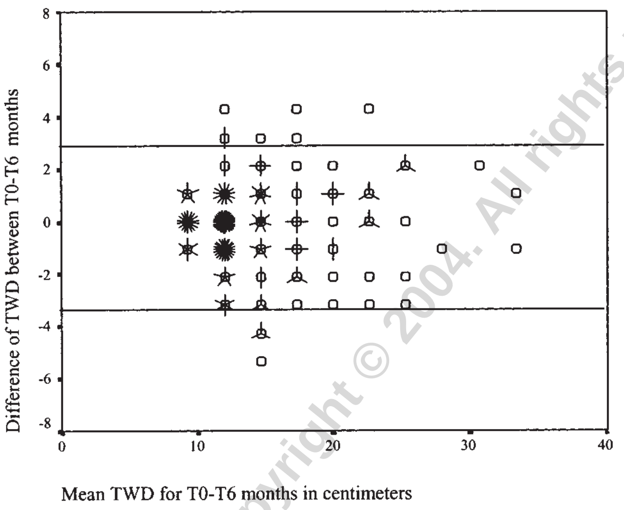

Figure 4. Graph of the difference scores against the mean score for TWD in cm in all patients, each spike of the sunflower representing one case.

Analyzing the 95% limits of agreement of OWD and TWD in all patients compared to kyphosed patients shows that the 95% limits of agreement are comparable for OWD and TWD (Table 4). However, the 95% limits of agreement in kyphosed patients is increased compared to all patients: for OWD –4.8 and 4.2 cm and for TWD –4.2 and 3.6 cm.

The mean change in BASDAI between T_0 and T_6 on a group level was 0.09 (SD 1.76). In 100 patients the BASDAI was stable between T_0 and T_6. In this group of patients the limits of agreement for OWD were –3.2 and 3.2 cm and for TWD –3.3 and 3.3 cm. Examining only kyphosed patients with a stable BASDAI (n = 56), the limits of agreement were –4.4 and 4.4 cm for OWD and –4.5 and 4.5 for TWD. These limits of agreement do not differ from the limits of agreement found in analysis of the whole group. It thus seems justified in this situation to use measurements with a time interval of 6 months for test-retest purposes. Furthermore, these data suggest there was no systematic error.

Of the 217 patients, 115 (53%) were not able to touch the wall with their occiput at both T_0 and T_6. These patients were classified as kyphosed. Table 4 shows the limits of agreement comparing baseline and T_6 for OWD and TWD calculated for these patients. The results are presented for the whole range of OWD and TWD and divided into quartiles. No consistent pattern over the quartiles was detected for all time intervals. We concluded that the limits of agreement are comparable over the whole scale and thus a more severe kyphosis does not evoke more measurement error. For comparison of OWD and TWD it is appropriate to use the limits of agreement not divided in quartiles of the scale.

DISCUSSION

In measurement of both OWD and TWD it is required that the chin is "in the neutral position." We hypothesized that

Table 2. Summary statistics (217 patients) of the average scores of occiput-to-wall distance (OWD) and tragus-to-wall distance (TWD) with a 6-month interval, summary statistics of the difference scores of OWD and TWD with a 6-month interval, and 95% limits of agreement of scores of OWD and TWD with a 6-month interval.

Paired Data	Average of the 2 Observations, cm* Mean (SD)	Median (min; max)	Difference of Paired Observations, cm** Mean (SD)†	95% Limits of Agreement of Paired Observations, cm
OWD T_0–T_6	3.8 (5.5)	1.2 (0.0; 24.9)	–0.2 (1.7)	–3.4; 3.1
TWD T_0–T_6	14.2 (4.7)	12.2 (8.2; 34.0)	–0.2 (1.6)	–3.4; 2.9
OWD T_6–T_{12}	3.8 (5.6)	1.4 (0.0; 25.9)	0.6 (1.8)	–2.9; 4.2
TWD T_6–T_{12}	14.0 (4.8)	12.3 (7.5; 34.0)	0.3 (1.8)	–3.3; 3.9
OWD T_{12}–T_{18}	4.1 (5.7)	2.0 (0.0; 28.2)	0.2 (1.7)	–3.1; 3.5
TWD T_{12}–T_{18}	14.3 (5.1)	12.6 (6.5; 35.0)	0.4 (1.8)	–3.1; 3.9
OWD T_{18}–T_{24}	4.4 (5.9)	2.5 (0.0; 29.5)	0.4 (1.5)	–2.5; 3.3
TWD T_{18}–T_{24}	14.6 (5.4)	12.9 (6.4; 39.0)	0.3 (1.7)	–3.1; 3.7

* Average score of the 2 observations = (observation 1 + observation 2)/2. ** Difference score between the 2 observations = observation 1 minus observation 2. † Standard deviation of the difference scores (i.e., the $SD_{difference}$), which is an estimate of random measurement error used to calculate the 95% limits of agreement.

Table 3. 95% limits of agreement for the whole group of patients and 95% limits of agreement for the group of patients with OWD > 0.

Paired Data	95% Limits of Agreement of Paired Observations in All Patients, cm N = 217	95% Limits of Agreement of Paired Observations in Kyphosed Patients, cm N = 115
OWD T_0–T_6	–3.4; 3.1	–4.8; 4.2
TWD T_0–T_6	–3.4; 2.9	–4.2; 3.6
OWD T_6–T_{12}	–2.9; 4.2	–3.7; 5.3
TWD T_6–T_{12}	–3.3; 3.9	–3.3; 4.5
OWD T_{12}–T_{18}	–3.1; 3.5	–3.9; 4.3
TWD T_{12}–T_{18}	–3.1; 3.9	–3.5; 4.3
OWD T_{18}–T_{24}	–2.5; 3.3	–4.3; 3.1
TWD T_{18}–T_{24}	–3.1; 3.7	–3.5; 4.7

involuntary flexion or (hyper)extension of the head, which takes place in the atlanto-occipital joint and the cervical spinal joints and can occur together with thoracic spine extension, would influence the OWD more than the TWD because, compared to the occiput, the tragus lies closer to the sagittal axis, about which the atlanto-occipital joint flexion and extension occur. In other words, patient variance would influence the OWD more compared to TWD and thus the OWD would be more prone to measurement error. On the other hand, variation in rotation of the head could influence the TWD more compared to the OWD. However, analyzing our data we found similar ICC and similar limits of agreement for OWD and TWD, which did not confirm a differential aspect in measurement error between the 2 instruments.

We did not perform a formal test-retest method in assessing the measurement error of the OWD and TWD with a short time interval. However, in our view it is appropriate to use the data with a 6-month interval in a disease with very slow progression. The various analyses we performed do underscore this assumption.

The OWD and TWD instruments did not behave differently across different ranges of the measurement. Auléley, *et al*[9] reported in an international study on 120 patients with AS a 95% limit of agreement between –2.8 and 2.9 cm for OWD, which is a little lower than in our group. In their investigation TWD was not studied. However, we do not know the percentage of patients with a normal OWD in that study; if this was higher this may explain the small difference, as we showed that measurement error is increased in kyphosed patients. In calculation of the BASMI[4,5], an intuitively created index based on the literature, the TWD and not the OWD was used. As stated, this measurement was among 20 separate measurements historically used in the assessment of AS. Pile, *et al*[10] found an excellent interobserver reliability, represented by an interobserver coefficient of reliability of 0.97 for TWD. No interobserver coefficient of reliability was given for OWD. Viitanen, *et al*[11] reported an interobserver ICC of 0.94 for both OWD and TWD and an intraobserver ICC of 0.98 for OWD and 0.90 for TWD.

Table 4. Ranges of OWD and TWD of patients with OWD > 0 at T_0 or T_6 are given, with 95% limits of agreement between T_0 and T_6 for the whole group of patients with OWD > 0 divided in quartiles.

	Whole Range	T_0–T_6, N = 115 1st Quartile, n = 29	2nd Quartile, n = 29	3rd Quartile, n = 29	4th Quartile, n = 28
OWD T_0–T_6					
Range	0.4–24.9	0.4–2.6	2.7–5.3	5.4–9.8	10.0–24.9
95% limits of agreement, T_0–T_6	–4.8; 4.2	–5.4; 4.3	–5.3; 4.5	–4.4; 3.8	–4.2; 3.7
TWD, T_0–T_6					
Range	10.0–34.0	10.0–15.5	11.8–17.1	13.1–20.6	16.6–34.0
95% limits of agreement, T_0–T_6	–4.2; 3.6	–4.8; 3.0	–4.9; 4.1	–3.3; 3.7	–4.2; 4.0

But these investigators did not present 95% levels of agreement, and a high ICC does not rule out an unacceptable level of measurement error[12].

Bellamy, *et al* reported the median of the minimally clinically important difference estimates (MCIDE; as defined by an expert panel with use of Delphi rounds) for TWD as 3.0 cm and for OWD as 1.25 cm[13]. In both our study and the study of Auléley, *et al*[9] the 95% limits of agreement are beyond the reported MCIDE from this expert panel, which suggests that the MCIDE cannot reliably be distinguished from measurement error.

Our conclusion is that we found no difference in reliability between OWD and TWD. Although we feel the TWD is in general easier to perform compared to OWD in patients with AS, we recommend the OWD measurement over the TWD — that in OWD measurement, a value of zero easily distinguishes patients with normal thoracic spine extension from kyphosed patients. This argument is especially valid in the research setting, where a value of 0 for OWD clearly means no kyphosis and a value, for example, of 12 for TWD could mean no kyphosis or beginning kyphosis. We think this outweighs the minor disadvantages of the feasibility of measurement of OWD.

REFERENCES

1. Bellamy N. Musculoskeletal clinical metrology. Dordrecht: Kluwer Academic Publishers; 1993.
2. van der Heijde D, Calin A, Dougados M, Khan MA, van der Linden S, Bellamy N. Selection of instruments in the core set for DC-ART, SMARD, physical therapy, and clinical record keeping in ankylosing spondylitis. Progress report of the ASAS Working Group. J Rheumatol 1999;26:951-4.
3. Bellamy N. Clinimetric concepts in outcome assessment: The OMERACT filter. J Rheumatol 1999;26:948-50.
4. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol 1994;21:1694-8.
5. Jones SD, Porter J, Garrett SL, Kennedy LG, Whitelock H, Calin A. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI) [letter]. J Rheumatol 1995;22:1609.
6. Spoorenberg A, van der Heijde D, de Klerk E, et al. A comparative study of the usefulness of the Bath Ankylosing Spondylitis Functional Index and the Dougados Functional Index in the assessment of ankylosing spondylitis. J Rheumatol 1999;26:961-5.
7. van der Heijde DM, van 't Hof MA, van Riel PL, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916-20.
8. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
9. Auléley GR, Benbouazza K, Spoorenberg A, et al. Evaluation of the smallest detectable difference in outcome or process variables in ankylosing spondylitis. Arthritis Rheum 2002;47:582-7.
10. Pile KD, Laurent MR, Salmond CE, Best MJ, Pyle EA, Moloney RO. Clinical assessment of ankylosing spondylitis: a study of observer variation in spinal measurements. Br J Rheumatol 1991;30:29-34.
11. Viitanen JV, Heikkila S, Kokko ML, Kautiainen H. Clinical assessment of spinal mobility measurements in ankylosing spondylitis: a compact set for follow-up and trials? Clin Rheumatol 2000;19:131-7.
12. Atkinson G, Nevill A. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-38.
13. Bellamy N, Buchanan WW, Esdaile JM, et al. Ankylosing spondylitis antirheumatic drug trials. III. Setting the delta for clinical trials of antirheumatic drugs — results of a consensus development (Delphi) exercise. J Rheumatol 1991;18:1716-22.