Validation of the Turkish version of the visual analog scale spine score in patients with spinal fractures

Osman YARAY¹, Burak AKESEN¹, Gökhan OCAKOĞLU², Ufuk AYDINLI¹

¹Department of Orthopaedics and Traumatology, Faculty of Medicine, Uludağ University, Bursa, Turkey; ²Department of Biostatistics, Faculty of Medicine, Uludağ University, Bursa, Turkey

Objective: The visual analog scale spine score (VASSS) is a valid and reliable instrument for outcome assessment of patients with thoracic and lumbar spine fractures. The aim of this study was to prepare a Turkish version of the VASSS and to validate its use for assessing treatment outcomes in Turkish patients with spinal trauma.

Methods: The German version of the VASSS was blindly and independently translated into Turkish by three translators and modified by a team. Fifty patients who had been surgically treated for thoracic or lumbar fracture and a group of 50 healthy controls were evaluated using the VASSS, Oswestry Disability Index (ODI), Roland-Morris Disability Questionnaire (RMDQ), and Short Form 36 (SF-36). The Cronbach’s alpha was performed to test the internal consistency of the score.

Results: The Cronbach’s alpha coefficient was calculated as 0.965 in the overall assessment of the scale. Criterion validity measured by comparing the VASSS responses with the results of ODI, RMDQ, and SF-36 physical component (for ODI r=0.881, p<0.001; for RMDQ r=0.882, p<0.001; for SF-36 r=0.824, p<0.001). Construct validity tested by factor analysis yielded a factorial structure of the questionnaire with 64.7% of cumulative percentage of explained variance, and Turkish version of the VASSS showed a similar structure than the original version.

Conclusion: The Turkish version of the VASSS is a reliable and valid instrument to assess the outcome in patients with thoracic or lumbar spinal fractures in the Turkish population.

Key words: Fracture; Oswestry Disability Index; spine score; Turkish; visual analog scale spine score.

Traumatic vertebral fractures and fracture-dislocations are severe injuries seen especially in young individuals. Although spinal injuries constitute 6% of all traumatic injuries, their impact on patients’ socioeconomic status is greater.⁴⁻⁶ The restoration of normal function is the main objective in the treatment of thoracic and lumbar spinal fractures. The results of thoracic and lumbar spinal fracture treatment was once evaluated in terms of range of motion, muscle strength and radiological results of treatment and referred to by some authors as a “surrogate” outcome.⁷ In recent years, however, more emphasis has been placed on the measurement of symptoms, functional status, satisfaction with treatment, and health care cost associated with spinal interventions.⁸⁻¹⁰ There has also been a growing recognition that patients’ perspectives are essential, both in making medical decisions and in judging the results of treatment.

The most commonly used and well-known disability scales for patients with back pain are the Oswestry Disability Index (ODI) and Roland-Morris
Disability Questionnaire (RMDQ). The Short Form 36 (SF-36) is the most commonly used generic test. Although the SF-36 is valid for measuring morbidity and surgical outcomes in common spinal disorders, it is not specific to any disease or condition.\(^\text{11}\)

The visual analog scale (VAS), developed by Freyd, is a well-known measurement tool for pain, consisting of a single 100 mm line.\(^\text{12-15}\) The VAS Spine Score (VASSS) was developed by Knop et al. (Work Group Spine, German Trauma Association [DGU]) in 2001.\(^\text{16}\) It consists of 19 questions which are scored on a visual analog scale. With the VASSS, the patient’s perception of pain and restriction in activities, related to back problems, can be measured. The score is calculated by taking the average scores of all answered questions and can be any value between zero (severe disability) and 100 (no disability). The most negative responses are on the left, the most positive responses are on the right side of each VAS line. In order to facilitate the answers, descriptions such as “always, for hours, constantly” or “never, rarely, too short” were added to the corresponding ends of the scales. In addition, small graphics of happy face/smiling sun on the right end and unhappy face/rain clouds on the left end are displayed. In the original study, most of the patients did not answer question 7 “Wie gut wirken die Schmerzmittel dann?” (How good are the painkillers?), since they did not use any medication for pain. Therefore, the developers of the VASSS suggested the exclusion of question 7, which we followed in our study.

The VASSS is originally in German. It is clear that a scale cannot be transferred directly from one culture to another without being reevaluated for the new conditions.\(^\text{17}\) Therefore, a simple direct translation of a questionnaire does not permit its use in clinical trials. The translation must be validated to obtain an equivalent questionnaire and to allow the comparability of data.

To date, a Turkish version of the VASSS has not been validated. The aim of this study was to translate and culturally adapt the Turkish version of the VASSS and to validate its use for assessing the outcome of patients with thoracic and lumbar spinal fractures.

**Patients and methods**

Patients operated on for a fracture of the thoracic or lumbar spine by the same surgeon in our institution between 1995 and 2005 were included in the study. Minimum follow-up time was 4 years. Patients with a pathological fracture, neurological deficit or insufficient command of the Turkish language were not included in the study. Fifty patients meeting the above mentioned criteria were included in the study. In addition, a group of 50 healthy persons were included in the study as a control group. These subjects were recruited from the hospital staff, were in the same age group and declared no history of any spinal surgery or any diagnosed chronic disease. The Medical Ethics Committee of Uludağ University approved the study protocol.

For the translation process, we used the recent guidelines for cross-cultural adaptation.\(^\text{18,19}\) Three translations from German to Turkish were performed by three different and independent translators whose native language was Turkish. One of the translators was aware of the process purpose and the concepts involved in the instrument, in order to obtain a better idiomatic and conceptual translation, rather than a literal one, and to render the intended measurement more reliable. The other two translators were unaware of the translation objective, which was useful in eliciting unexpected meanings from the original tool. The translations were then retranslated into German, compared with the original German VASSS and checked for inconsistencies.

The Turkish versions were reviewed by a team including three translators, three orthopedic surgeons, one physiotherapist, and one Turkish language teacher to assess the necessity of performing a cultural adaptation. They decided to change the statement “trenle yolculuk etmek” (traveling by train) to “otobüsle yolculuk etmek” (traveling by bus) in item 16, as bus travel is more common in Turkey. The final stage of adaptation was to test the pre-final version. Fifteen people were tested in this stage. The statements “fiziksel aktiviteler” (physical activities) in item 4 and 5 were not clearly understood by all patients in this stage, so the descriptions “gün içinde yapılan hareketler” (daily basis activities, e.g. climbing stairs) were added instead. This was finalized after slight changes.

Two common forms of reliability are test-retest reliability and internal consistency. Test-retest reliability measures the stability over time by administering the same test to the same subjects at two points in time. In our study population, most patients were from rural parts of the city and most did not return to the hospital for retesting. Therefore, we could not achieve the test-retest reliability.
The internal consistency of a scale relates to its homogeneity. The coefficient of internal consistency is mainly assessed with Cronbach’s alpha. It is suggested that the value of alpha should be above 0.80 for acceptance as high internal consistency.\[20\] The internal consistency reliability of the VASSS was assessed by calculating “if item deleted” using Cronbach’s alpha and “item–total correlation” coefficient for each item of the questionnaire.

Criterion validity was measured by comparing the VASSS responses with other measurements performed at the same time. For this purpose all participants completed the Turkish version of VASSS, the Turkish validated version of ODI,\[21\] the Turkish validated version of RMDQ,\[22\] and the Turkish validated version of SF-36.\[23\] Criterion validity was measured by the Spearman’s correlation coefficient. The coefficients were accepted as follows: 0.81-1.0 “excellent”, 0.61-0.80 “very good”, 0.41-0.60 “good”, 0.21-0.40 “fair”, and 0-0.20 “poor”.\[24,25\] Construct validity was assessed by explanatory factor analysis.

The ODI is a disease-specific instrument for the assessment of the affect of activity on pain intensity, consisting of a 10-item ordinal scale instrument. The total score ranges from 0 to 100, where 100 is the worst disability. The RMDQ is a validated questionnaire to measure disability due to back pain. It consists of 24 items with “yes” or “no” answers. The score could, thus, vary from zero (no disability) to 24 (severe disability). The SF-36 scale contains eight sub-scales: physical functioning, role restriction due to physical problems, pain, general perception of health, social functioning, role restriction due to emotional problems, mental health, and vitality. The first four sub-scales are physical components and the later four are mental components. Scores vary from 0 to 100; higher scores indicate better results.

In the present study, the mean of the first four sub-scales are calculated to find the physical component (SF-36 PC) score. For statistical evaluation, the scores of RMDQ and ODI were transformed to a percentage by the following formulas: (1 - (n/24) x 100) and (100 - n), respectively. This resulted in a score of 0 when the RMDQ was 24 and the ODI was 100, and a score of 100 when the RMDQ and ODI were 0, indicating no disability at all. No transformation was needed for SF-36 PC. Continuous variables were represented as mean ± standard deviation. A value of p<0.05 was considered statistically significant. All data were analyzed with SPSS version 13.0.

### Results
In the patient group, a total number of 50 patients with a mean age of 48±14.25 (range: 23 to 77) years were treated either because of thoracic or lumbar spinal fracture with ventral, dorsal or combined fusion surgery. Mean follow-up time after surgery was 9±3.09 (range: 4 to 15) years. The control group consisted of 50 people with a mean age of 48±8.25 (range: 34 to 65) years. The mean scores of instruments for both patient and control groups are summarized in Table 1.

<table>
<thead>
<tr>
<th>Item number</th>
<th>Item-total correlation</th>
<th>Cronbach’s alpha “if item deleted”</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.795</td>
<td>0.963</td>
</tr>
<tr>
<td>2</td>
<td>0.715</td>
<td>0.964</td>
</tr>
<tr>
<td>3</td>
<td>0.715</td>
<td>0.964</td>
</tr>
<tr>
<td>4</td>
<td>0.809</td>
<td>0.962</td>
</tr>
<tr>
<td>5</td>
<td>0.804</td>
<td>0.962</td>
</tr>
<tr>
<td>6</td>
<td>0.711</td>
<td>0.964</td>
</tr>
<tr>
<td>7</td>
<td>0.733</td>
<td>0.963</td>
</tr>
<tr>
<td>8</td>
<td>0.793</td>
<td>0.963</td>
</tr>
<tr>
<td>9</td>
<td>0.882</td>
<td>0.961</td>
</tr>
<tr>
<td>10</td>
<td>0.779</td>
<td>0.963</td>
</tr>
<tr>
<td>11</td>
<td>0.860</td>
<td>0.961</td>
</tr>
<tr>
<td>12</td>
<td>0.752</td>
<td>0.963</td>
</tr>
<tr>
<td>13</td>
<td>0.721</td>
<td>0.964</td>
</tr>
<tr>
<td>14</td>
<td>0.818</td>
<td>0.963</td>
</tr>
<tr>
<td>15</td>
<td>0.764</td>
<td>0.963</td>
</tr>
<tr>
<td>16</td>
<td>0.698</td>
<td>0.964</td>
</tr>
<tr>
<td>17</td>
<td>0.790</td>
<td>0.963</td>
</tr>
<tr>
<td>18</td>
<td>0.834</td>
<td>0.962</td>
</tr>
</tbody>
</table>

### Table 1. Demographic data and test results of patient and control groups.

<table>
<thead>
<tr>
<th></th>
<th>Patient group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Mean age</td>
<td>48±14.25</td>
<td>48±8.25</td>
</tr>
<tr>
<td>Male</td>
<td>34(68%)</td>
<td>25(50%)</td>
</tr>
<tr>
<td>Female</td>
<td>16(32%)</td>
<td>25(50%)</td>
</tr>
<tr>
<td>VASSS</td>
<td>69.6±21.80</td>
<td>85.9±19.58</td>
</tr>
<tr>
<td>ODI</td>
<td>78.6±19.30</td>
<td>93.4±9.14</td>
</tr>
<tr>
<td>RMDQ</td>
<td>66.4±30.47</td>
<td>86.0±21.72</td>
</tr>
<tr>
<td>SF-36 PC</td>
<td>63.3±23.60</td>
<td>76.2±18.88</td>
</tr>
</tbody>
</table>

Internal consistency of the VASSS for individual items is shown in Table 2. Cronbach’s alpha coefficient was calculated 0.965 by the overall assessment
of the scale. The item–total correlation was found to be greater than 0.25.

Criterion validity was tested by determining the correlation between VASSS and ODI, RMDQ, and SF-36 PC. The resulting correlations were excellent (for ODI r=0.881, p<0.001; for RMDQ r=0.882, p<0.001; for SF-36 PC r=0.824, p<0.001) as shown in Table 3. The VASSS results of the patient and control groups were significantly different (Mann-Whitney U test, p<0.001) (Table 4). The VASSS was able to discriminate between the patient group and control group.

**Table 3.** Criterion validity of VASSS against ODI, RMDQ and SF–36 PC.

<table>
<thead>
<tr>
<th></th>
<th>VASSS</th>
<th>ODI</th>
<th>RMDQ</th>
<th>SF-36 PC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p</td>
<td>r</td>
<td>p</td>
</tr>
<tr>
<td>ODI</td>
<td>0.881</td>
<td>&lt;0.001</td>
<td>0.882</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RMDQ</td>
<td>0.882</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF–36 PC</td>
<td>0.824</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.** Comparison of mean VASSS scores between patient and control group.

<table>
<thead>
<tr>
<th></th>
<th>N (number)</th>
<th>Mean</th>
<th>SD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient group</td>
<td>50</td>
<td>69.60</td>
<td>21.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Control group</td>
<td>50</td>
<td>85.94</td>
<td>19.58</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

We followed the well-documented sequential process of adaptation of outcome measure for use in different cultures. The reliability and validity of the scale obtained from the current study were considered sufficient for this research. Reliability checks of the Turkish version of the VASSS proved that it is a dependable scale. We tested the reliability of our scale using Cronbach’s alpha coefficient. As most patients refused the retest, the reliability of the scale was determined by applying the scale once. In the event that the scale is applied once, the reliability of internal consistency is examined and the coefficient ranges between 0 and 1. In this study, reliability results were considered to be excellent for all items and values comparable to what was obtained from the original version. The Cronbach’s alpha was 0.965 in the present study and 0.916 by Knop et al. in the original version.

The criterion validity assesses the scale efficiency. In the analysis of criterion validity, there were strong correlations between the VASSS and the ODI, RMDQ, and SF-36. Our results were in agreement with those previously reported. Siebenga et al.
found the correlation of the VASSS and RMDQ to be 0.870 \((p<0.001)\) in the operatively treated group.\[^{28}\] The correlation between VASSS and RMDQ was 0.850, and 0.870 between the VASSS and the SF-36 physical functioning, in the study of the functional outcome of type A spinal fractures.\[^{29}\]

Construct validity is used to investigate to what degree any particular measure relates to other measures in accordance with the hypothesis on the measured parameters.\[^{30}\] Internal consistency also reveals the construct validity.\[^{31}\] To assess construct validity, 18 items were factor-analyzed using the method of principal components analysis, with the Quartimax rotation as the orthogonal solution. The results of the exploratory factor analysis showed that one factor was clearly associated with all items and yielded a factorial structure of the questionnaire with a 64.7% cumulative percentage of explained variance. The Turkish version of the VASSS showed a similar structure to the original version.

The most commonly used back-specific measures are the RMDQ and the ODI. The RMDQ measures 24 activity limitations due to back pain with 2 response options (“yes” or “no”). The ODI consists of 10 items assessing the level of pain interference with physical activities with 6-level response options. Completing the RMDQ, subjects have to agree or disagree with the statements. Upon completion of the ODI, patients have to choose one of six choices to determine their disability. In the event that VASSS subjects answer the questions on a 100 mm VAS, using a mark (tick) and determine their restriction or pain level by their own judgment, VASSS is 100% subjective.

The RMDQ and ODI completion time is approximately 5 minutes.\[^{32}\] Once the subject clearly understands how to answer a VAS or is familiar with a VAS, the VASSS completion time is about 3 minutes. However, manual scoring of the VASSS is much longer than the other two instruments. Therefore, Knop et al. developed a computerized analysis system, which lowered the VASSS scoring time to approximately 20 seconds.

Our study had some limitations, including the small number of patients and lack of test-retest validity. However, in the original study of Knop et al. 53 patients were included in the patient group.

We think that a study based on the Turkish population with a larger number of patients would increase the value of our present study. Additionally, in further studies, the validation of the Turkish VASSS should be tested in different spinal conditions.

In conclusion, the Turkish version of the VASSS is a reliable and valid measurement for the assessment of treatment outcomes in patients with thoracic or lumbar spinal fractures in the Turkish population.

**Acknowledgement**

The authors thank Dr. Christian KNOP, for his permission to translate the VASSS into Turkish, and the members of the committee (Ali ÖZER, Nevzat ABACIOĞLU, Murat MORAY, Remzi ÖZERDEMOĞLU, translators and Özlem YARAY, Turkish language teacher) for their cooperation.

**Appendix**

**Turkish version of the VASSS:**

1. Bel ağrısı nedeniyle uygunuz ne sıklıkla bölünüyor?
2. Dinlenme sırasında ne sıklıkla bel ağrınızı oluyor?
3. Dinlenme sırasında oluşan bel ağrınızın şiddeti nedir?
4. Fiziksel aktivite (gün içinde yapılan hareketler) sırasında ne sıklıkla bel ağrınız oluyor?
5. Fiziksel aktivite (gün içinde yapılan hareketler) sırasında oluşan bel ağrınızın şiddeti nedir?
6. Bel ağrınız olduğu zaman ne sıklıkla ağrı kesici kullanıyormusuz?
7. Bel ağrınız olmadan ne kadar süre oturabilirisiniz?
8. Bel ağrınız öne doğru eğilmeniz ne kadar engelliyor (örneğin buluşık yikarken)?
9. Bel ağrınız kişinin, mesleğinizi yapmanızı ne kadar engelliyor?
10. Bel ağrınız bir şey kaldırmanızı ne kadar kıstlıyor?
11. Bel ağrınız ev işleri yapmanızı ne kadar kıstlıyor?
12. Bel ağrınız olmadan ne kadar süre ayakta durabilirsiniz?
13. Bel ağrınız olmadan ne kadar süre yürüyebilirsiniz?
14. Bel ağrınız koşmanızı ne kadar engelliyor?
15. Bel ağrınız günlük işlerinizi ne kadar engelliyor (yemek yeme, banyo yapma gibi)?
16. Bel ağrınız olmadan ne kadar süre yolculuk yapabilirsiniz (araba sürmek, otobüsle yolculuk gibi)?
17. Bel ağrınız cinsel hayatınızı ne kadar kısıtlıyor?
18. Bel ağrınız bir eşya veya yük kaldırmanızı ne kadar kısıtlıyor?

**Conflicts of Interest:** No conflicts declared.
References