


Reliability and Validity of the German Version of the AO Spine Patient Reported Outcome Spine Trauma Questionnaire

Global Spine Journal
2023, Vol. 0(0) 1–7
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/21925682231156124
journals.sagepub.com/home/gsj


Sonja Häckel, MD^{1,†} , Katharina A. C. Oswald, MD^{1,†} , Leonie Koller¹, Lorin M. Benneker², Layla A. Benneker³, Said Sadiqi, MD, PhD⁴, F. Cumhur Oner, MD, PhD⁴, and Moritz C. Deml, MD¹ 

Abstract

Study Design: A single-center validation study.

Objective: To translate and cross-culturally adapt the AO Spine PROST (Patient Reported Outcome Spine Trauma) into German, and to test its psychometric properties among German-speaking Swiss spine trauma patients.

Methods: Patients were recruited from a level-I Swiss trauma center. Next to the AO Spine PROST, the EQ-5D-3L questionnaire was used for concurrent validity. Questionnaires were filled out at two-time points for test-retest reliability. Patient characteristics were analyzed using descriptive statistics. For content validity, floor, and ceiling effects, as well as any irrelevant and missing questions were analyzed. Construct validity of the AO Spine PROST questionnaire to the EQ-5D-3L was tested using Spearman correlation tests.

Results: The AOSpine PROST was translated and adapted into German using established guidelines. We included 179 patients. The floor effect for all items was well within the optimal range (below 15%), while the ceiling effect of seven items was within the optimal range. None of the items displayed a problematic floor or ceiling effect. The overall test-retest reliability of the total PROST score was excellent, with an ICC of .83 (95% CI .69-.91). The Spearman correlation coefficient between the total PROST summary score and EQ-5D-3 L was $\rho = .63$.

Conclusions: The German version of the AO Spine PROST questionnaire demonstrated very good validity and reliability results.

Keywords

outcome instrument, AO spine PROST, patient perspective, spine trauma

Introduction

Measuring the individuals' health-related quality of life (HRQoL) after spine trauma is highly relevant. These injuries - often caused by high-velocity accidents, with the thoracolumbar spine (TL) being the most affected region - can lead to severe impairments of the patients and disabilities. From a social and economic point of view, this leads to a global health concern.^{1,2} Specific outcome measurements provide the opportunity to compare outcomes of different treatment strategies in a certain patient population. The results can contribute

¹ Department of Orthopaedic Surgery and Traumatology, Inselspital, University Hospital Bern, University Bern, Bern, Switzerland

² Ortho-Spine, Sonnenhof Spital Bern, Bern, Switzerland

³ Department for Mathematics and Statistics, University of Bern, Bern, Switzerland

⁴ Department of Orthopaedics, University Medical Center Utrecht, Utrecht, the Netherlands

[†]Both authors contributed equally to this publication.

Corresponding Author:

Sonja Häckel, Department of Orthopaedic Surgery and Traumatology, Inselspital, University Hospital Bern, University Bern, Freiburgstrasse 18, Bern 3010, Switzerland.
Email: sonja.haekkel@insel.ch



Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

to the development and adaptation of patient-specific treatment plans. However, until the development of the spine-specific AO Spine Patient Reported Outcome Spine Trauma (PROST) questionnaire, only non-specific questionnaires were used. The questionnaires that are among the most used to assess overall health include the EQ-5D instruments and the Short Form 36 (SF-36) questionnaire. In addition, there are numerous other questionnaires for assessing patient-reported physical disability due to low back pain, such as the Roland-Morris Disability Questionnaire or the Oswestry Disability Index. However, these instruments were developed to assess the quality of life of patients with chronic low back pain, degenerative disease, or disability in low back pain. Their use as an instrument for clinical outcomes, quality of life, and functional limitations due to spine trauma is limited.³⁻⁶

In 2017, the AO Spine Knowledge Forum (KF) Trauma developed a disease-specific outcome measure for patients suffering from spine trauma: The PROST questionnaire, initially developed in Dutch. This questionnaire contains 19 items related to different areas of functional aspects in everyday life.⁷ Patients with spinal injuries leading to complete paralysis or suffering from multiple traumatic injuries (Injury Severity Score >15) were excluded from the developmental process of the AO Spine PROST. This aimed to identify limited life quality solely related to spinal injury and exclude the overwhelming factors of paralysis and polytrauma. The recent validation studies of the Dutch, English, and Nepali versions of the PROST showed very good results.⁷⁻⁹ However, a psychometrically validated outcome measure for analyzing the trauma-related quality of life limitations and treatment success of German-speaking patients with spinal cord injuries is missing so far.⁶

This study aimed to (1) translate the AO Spine PROST into German, including a cross-cultural adaption, and (2) test its psychometric properties in German-speaking patients suffering from an acute traumatic spinal injury.

Materials and Methods

Translation and Cross-Cultural Adaptation

Translation and cross-cultural adaption were performed using established guidelines.¹⁰ Due to the missing analysis of clinical data, the institutional review board decided no explicit ethics approval was necessary; a waiver was granted. All patients signed the informed consent to be included in the study. The Dutch version of the AO Spine PROST was used and translated into German according to the method described earlier^{7,10}: (1) Forward translations, (2) synthesis, (3) back translations, and (4) expert committee review, followed by validation.

The first stage comprised the initial translation. Two native bilingual German and Dutch speakers did the forward translation of the AO Spine PROST from Dutch into German independently. In stage 2, the two independent translators

summarized their findings into one synthesis. In stage 3, two separate, uninformed bilingual translators with Dutch as their native language did the back-translation from German to Dutch. Both were blinded to the concepts under investigation. In the next phase, all translations and written reports were reviewed by an expert committee to achieve equivalence between the four domains, semantics, idiomaticity, experience, and conceptualism, developing the final German version of the AO Spine PROST questionnaire. In the fifth stage, the preliminary version was tested on 30 patients treated in our spine trauma outpatient clinic. Patients were asked to fill out the EQ-5D-3L questionnaire. Additionally, they were questioned whether and how they would change or improve the questionnaire. This pre-testing provided information on how patients interpreted the items of the questionnaire and ensured that the translated version retained its equivalence in an administered situation. After the validation process, no further changes were performed to the German version of the AO Spine PROST.

Patients and Study Procedures

All patients in our clinic (level-1 trauma center) were prospectively screened for eligibility. Inclusion criteria were history of spinal column injury as the primary diagnosis, age >18 years, and capability to understand and adequately fill out the questionnaires. The inclusion and exclusion criteria were in line with those in the original AO Spine PROST. Exclusion criteria were multiple injuries (Injury Severity Score [ISS] > 15) or an American Spinal Cord Injury Association [ASIA] impairment grade A or B at discharge or transfer from the hospital.

Instruments

The questionnaires were given to the patients in paper form. Each item of the PROST questionnaire is accompanied by a numerical rating scale ranging from 0 to 100. Here, 0 indicates no function at all, and 100 is the level of function before the injury. Smileys support this scale at each end of the ruler (supplementary material). For validity purposes, the AO Spine PROST was compared with the EQ-5D-3L, which is a commonly used, generic outcome instrument for assessing HRQoL (scale from 0 to 1, where 1 indicates perfect health). The EQ-5D-3L consists of two pages, the EQ-5D description system and the EQ-5D visual analog scale (EQ VAS).^{11,12}

Statistical Analysis

The total PROST score was computed as the sum of all scores divided by the number of completed items, as previously described.⁸ It can range from 0 (worst) to 100 (best functional status). The responses to the EQ-5D-3L questionnaire items were combined and converted into a single utility estimate using the EU scoring algorithm and value set.¹³ At the same

time, the EQ-5D-3L was reported as pure patient-based ratings of the HRQoL.

Missing data were analyzed for each of the 19 PROST items and the total score was obtained by dividing the number of missing values by the number of respondents. Floor and ceiling effects were assessed by calculating the percentage of respondents who reported minimum and maximum possible scores, respectively, on individual PROST items and the overall score. If floor and ceiling effects occur in many patients, a questionnaire may not be able to detect deterioration or improvement in the patient's health status. Therefore, following established rules, we considered floor and ceiling effects larger than 70% to be problematic and effects smaller than 15% to be optimal.¹⁴

We also investigated the construct validity of the PROST questionnaire by assessing the degree to which the overall PROST scores relate to EQ-5D-3L utility scores. EQ-5D and PROST summary scores were not normally distributed, in which case it is best to use a non-parametric measure. Therefore, Spearman's Rho (ρ) corrected for ties was used to assess these correlations, whereby $\rho > .80$ was interpreted as an excellent outcome, .61-.80 very good, .41-.60 good, .21-.40 fair, and .00-.20 poor.¹⁵

Finally, we assessed the test-retest reliability of the PROST questionnaire to explore the questionnaire's stability over time by comparing PROST results at two points in time. We calculated intra-class correlation coefficients (ICC) and their 95% confidence intervals for each PROST item. ICC can occupy values between .0 and 1.0, with values of .6-.8 generally indicating good reliability and values above .8 indicating excellent test-retest reliability.¹⁶

Results

Cross-Cultural Adaptation Process

The first stage involved the initial translation, which was performed according to the guidelines developed by Beaton et al¹⁰ In the fifth step, the preliminary version was tested on 26 patients. This pretest provided information on how patients interpreted the items of each question. During the validation phase of the cross-culturally adapted and translated questionnaire, the first group of 26 patients was asked whether the questions were clearly worded and if the questions were difficult to understand. Only one patient added two comments. Question 3, recreation, and leisure (*Freizeit-Aktivitäten (Ausführen von Sport, Hobbies)*): patient's comment 'A more detailed formulation would be better here, eg 'I can perform these same leisure activities as before the accident'. Question 4, Social life (*Sozialleben (Pflegen von Kontakten mit Familie, Freunden und Bekannten)*): patient's comment 'Is it meant whether physical changes have an influence on the social life or also whether the accident has brought other changes?' However, these comments did not lead to an adaptation of the questionnaire, since it is explained in the questionnaire, that

Table 1. Patient Characteristics and Demographic Data.

Variable	Total (n = 179)
Total number of fractures	268
Age years ^a (SD)	48 (18)
Female, n (%)	68 (38)
Insurance class (%)	
1 (private)	15 (9)
2 (semi-private)	23 (13)
3 (public)	133 (78)
Cause of trauma (%)	
Sports/recreation	75 (42)
Fall	60 (34)
Motor/traffic vehicle accident	22 (12)
Other	22 (12)
AO fracture classification (%)	
Type A	95 (53)
Type B	65 (35)
Type C	18 (11)
Fracture level (%)	
Occipitocervical (C0-C2)	11 (6)
Subaxial cervical (C3-C7)	29 (16)
Thoracic (Th1-Th12)	63 (35)
Lumbar (L1-L5)	59 (33)
Sacral	17 (10)
Treatment (%)	
Surgical	128 (72)
Conservative	51 (28)
Surgical approach (%)	
Anterior	29 (23)
Posterior	77 (61)
Combine anterior-posterior	21 (16)
PROST ^a	76.0 (18.9)
EQ-5D (descriptive system) ^a	.73 (.20)
EQ VAS ^a	68.8 (19.4)

^aData are mean \pm standard deviation with percentages in parentheses.

patients should compare their 'function NOW to BEFORE the accident'.

Validation of the Cross-Cultural Adapted PROST Questionnaire

A total of 179 patients (≥ 18 years of age) from our Orthopedic Surgery and Traumatology department (level-1 trauma center) completed the PROST questionnaire between September 2017 and May 2021 after informed consent was given (Table 1). Eight patients were excluded due to dissent to the general consent. All patients had a history of spinal trauma and were either treated surgically or conservatively. The median interval between trauma and PROST questionnaire completion was 5.5 months (9 days to 8.7 years). To analyze the construct validity of the PROST tool, all patients completed the EQ-5D-3L questionnaire at the same time as the PROST questionnaire. In addition, the test-retest analysis of the PROST

questionnaire was assessed on a subsample of patients who completed the questionnaire a second time approximately one month later (median 28 days, range: 7 to 56 days).

Missing Values, Score Distribution, Floor, and Ceiling Effects

Table 2 shows that missing values were mainly observed in the items on work/study and sexual functioning. The floor effect for all items was well within the optimal range (below 15%), while the ceiling effect of seven items was within the optimal range. None of the items displayed a critical floor or ceiling effect. However, the reported scores were generally high, with the median of all 19 items greater than 60. The two items regarding urination and defecation had exceptionally high scores, with a ceiling effect exceeding 62%.

Construct Validity: Correlation Between PROST Summary Score and EQ-5D-3L

Spearman correlation coefficient between the PROST summary score and EQ-5D was $\rho = .63$, and the correlation between the PROST summary score and EQ VAS was $\rho = .69$, both $P < .001$. This is considered to be a very good outcome.

Test-Retest Reliability

Thirty-five patients completed the PROST questionnaire twice to test its stability over time. The overall test-retest reliability of the summary PROST score was excellent, with an ICC of .83 (95% CI .69-.91). None of the 19 PROST items showed

poor reliability. Three items showed excellent reliability, while thirteen items had good reliability (Table 3).

Discussion

This single-center validation study of the German version of the AO Spine PROST questionnaire investigated its psychometric properties after initial translation and cross-cultural adaptation. The German PROST questionnaire demonstrated very good construct validity and excellent test-retest reliability amongst German-speaking patients.

The PROST is a unique questionnaire and the first of its kind to focus specifically on spinal trauma patients.^{4,6} One of the unique features is that the patient is not asked only about his current condition as in many other HRQoL questionnaires, but to compare his current health status with that before the accident. PROST is designed to support trauma spine surgeons in the aftercare of their patients and to simplify spine research by allowing researchers to compare patient groups from multiple sites or across different regions of the globe. For the latter, the PROST must be available in many different languages. According to Beaton et al, the translation process by native speakers in Dutch and German guarantees a structured and well-adapted translation of the original Dutch version into German. The initial evaluation phase showed a very good understanding of all questions with only two comments from one patient which did not lead to any further changes to the questionnaire.

After cross-cultural adaptation was completed, we proceeded to test the German PROST in our patient cohort. During the analysis, we found that questions 2 (work/study 'Arbeit/Studium'

Table 2. Missing Data, Ceiling and Floor Effects of All Patients Included (n = 179).

AO Spine PROST Items	Missing (%)	Median (IQR)	Ceiling Effect: Max Score (%)	Floor Effect: Min Score (%)
1. Household activities	1.1	80 (60-95)	18.6	2.8
2. Work/study	30.2	75 (50-95)	0.0	6.2
3. Recreation and leisure	1.7	60 (30-80)	8.5	6.3
4. Social life	1.7	95 (70-100)	34.1	1.1
5. Walking	0.6	95 (74-100)	37.1	1.7
6. Travel	2.2	90 (70-100)	33.1	4.6
7. Changing posture	1.1	81 (60-95)	16.4	0.6
8. Maintaining posture	0.0	80 (60-90)	11.7	1.1
9. Lifting and carrying	0.6	66.5 (30-90)	11.8	3.4
10. Personal care	1.7	95 (80-100)	31.3	0.6
11. Urinating	1.1	100 (95-100)	62.2	1.1
12. Bowel movement	1.1	100 (95-100)	61.0	1.1
13. Sexual function	17.9	89 (60-100)	0.0	7.8
14. Emotional function	1.7	90 (80-100)	39.2	0.0
15. Energy level	1.1	81 (66-100)	26.0	0.0
16. Sleep	1.7	81.5 (70-100)	26.1	0.0
17. Stiffness of your neck and/or back	1.1	79 (60-90)	9.6	0.6
18. Loss of strength in your arms and/or legs	1.1	80 (64-95)	22.0	1.1
19. Back and/or neck pain	2.8	78.5 (50-90)	10.9	1.7

Table 3. Internal Consistency and Test-Retest Reliability Results. Internal Consistency is Sown for Each AOSpine PROST Item, Test-Retest Reliability was Assessed by Using ICC Along With its 95% CI.

AOSpine PROST Items	Median (IQR) 1st Test	Median (IQR) 2nd Test	ICC (95% CI)
1. Household activities	80 (60-90)	85 (60-95)	.73 (.53-.85)
2. Work/study	70 (35-80)	83 (50-95)	.61 (.25-.82)
3. Recreation and leisure	50.5 (24-80)	60 (50-83)	1.47 (.17-.69)
4. Social life	85 (50-99)	90 (75-100)	1.52 (.23-.72)
5. Walking	97 (90-100)	95 (80-100)	1.58 (.31-.76)
6. Travel	88 (60-100)	90 (70-100)	1.61 (.35-.78)
7. Changing posture	80 (60-90)	85 (71-95)	1.62 (.37-.79)
8. Maintaining body posture	80 (60-90)	80 (70-95)	1.66 (.43-.81)
9. Lifting and carrying	65 (30-80)	72 (40-89)	1.69 (.47-.83)
10. Personal care	90 (80-99)	95 (85-100)	1.63 (.37-.79)
11. Urinating	100 (95-100)	100 (80-100)	1.75 (.56-.86)
12. Bowel movement	100 (98-100)	100 (95-100)	1.74 (.55-.86)
13. Sexual function	90 (60-100)	90 (69-100)	1.96 (.92-.98)
14. Emotional function	90 (75-100)	90 (70-100)	.85 (.72-.92)
15. Energy level	87.5 (60-100)	85 (60-99)	1.82 (.67-.90)
16. Sleep	80 (60-97)	80 (70-100)	1.78 (.61-.88)
17. Stiffness of your neck and/or back	70 (50-85)	70 (50-90)	1.58 (.31-.76)
18. Loss of strength in your arms and/or legs	80 (66-90)	80 (60-91)	1.69 (.47-.83)
19. Back and/or neck pain	80 (50-90)	80 (50-90)	.67 (.44-.82)
PROST summary score	76.2 (61.6-87.8)	81.9 (68.2-93.8)	.83 (.69-.91)

and 13 (sexual function ‘*Geschlechtsverkehr*’) were not answered by many patients (30.2% and 17.9%, respectively). The PROST is designed in such a way that question 2 can be skipped if the patient did not work or study before the accident (this is also indicated in the questionnaire). In contrast, question 13 (sexual function) was skipped by many patients without being ‘permitted’ in the questionnaire (missing value of 17.9%). In this case, the total was calculated from 17 questions, even though this was not intended in the original version of the PROST.¹⁷ Another important factor determining the instrument’s acceptability (ie, completeness of questionnaires) is the way the information is collected. When the questionnaire is completed by patients on paper, a higher rate of missing values is to be expected, whereas this can be largely avoided with electronic data collection.¹⁸ This problem has not been discussed in the studies published so far; however, both the Dutch and the English validation of the questionnaires were performed with an online system, so this problem probably did not occur at all.

Furthermore, we analyzed the ceiling and floor effects of the German PROST. These can be observed when a significant number of patients reach either the maximum (ceiling) or minimum (floor) score. As a result, the true extent of their health status cannot be determined. The reported PROST scores were generally very high, suggesting that most patients included in the study had a quite mild disease or a very good outcome. In particular, the two items on urination and defecation had exceptionally high values. This is probably also because we had excluded ASIA-A, -B, and polytrauma patients. The feasibility of

using PROST in patients with complete spinal cord injury was investigated separately.¹⁹ As a result, the number of patients with a maximum score (ceiling effect) was high in most items, although not problematic in any of the items. This might also be related to the long time point after trauma. The floor effect for all items was well within the optimal range (below 15%). These results are comparable to the adapted version of the AO Spine PROST in English and the original Dutch version.^{7,17}

Since the PROST is the only validated questionnaire that specifically collects data from spine trauma patients, we needed another questionnaire for comparison. While other studies used the SF-36, we chose the EQ-5D-3L because it is one of the most common questionnaires for HRQoL and is routinely collected in our clinic.^{7,17} The Spearman correlation coefficient between the PROST sum score and the EQ-5D-3L was $\rho = .63$, demonstrating a consistent assessment of psychometric outcome analysis in spinal trauma patients.

This study has some limitations. The test-retest analysis was performed on all participants who completed the second questionnaire within 60 days of the first questionnaire (the median was 28 days). This is a relatively long interval (usually 2 weeks is targeted), which may have resulted in lower reliability compared with that observed with a shorter interval. However, the ICC of most items was reassuringly high despite the long interval. Another limitation of the study is the small sample size for the test-retest analysis ($n = 35$). Based on this sample, it was not possible to calculate the SEM and MDC 95% to determine the minimum detectable change. Larger

validation studies are needed for this type of analysis. Another limitation could be the long period between trauma and completion of the PROST questionnaire in some patients. In half of the patients, this period was 5.5 months or longer. Whether or not this is a problem depends mainly on the primary purpose of the PROST instrument in the study used. Our research shows that the PROST questionnaire provides reliable and consistent results for both long- and short-to-medium-term assessments after spinal trauma. However, the long-term reliability and validity could be confirmed in a recently published study.¹⁹ A final limitation is that we excluded patients with severe neurological impairment (AISA A and B) after spinal trauma. The original intention in developing the PROST questionnaire was to assess the outcomes of patients suffering from only mild to minor neurological impairments.^{6,17} However, the applicability of the PROST in patients with severe spinal cord injury was confirmed in a recently published study.¹⁹

In conclusion, the German version of the AO Spine PROST is a reliable and valid instrument for measuring HRQoL in German-speaking patients after spinal trauma, which can be used in clinical practice as well as in research.

The PROST questionnaires are available free of charge through the AO Spine website.²⁰

Acknowledgments

The authors thank Maïke Niewkamp for the questionnaire translation and Manuel Ottinger for his contribution to patient recruitment. We thank Lukas Staub of Talus Research Consulting (talus.com.au) for the methodological support, for performing the statistical analyses, and for helping with results reporting.

Author Contributions

MCD, LMB, SS, FCO, and SH set up the study protocol. KO, LK, LAB, and SH selected the data. Finally, SH, KO, LK, and MCD prepared the manuscript. All authors approved the final manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Informed Consent

This research does not include patient-specific data. Therefore, ethical consent was not necessary according to local regulations.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

ORCID iDs

Sonja Häckel  <https://orcid.org/0000-0002-9415-6633>

Katharina A. C. Oswald  <https://orcid.org/0000-0003-3713-9201>

Moritz C. Deml  <https://orcid.org/0000-0003-1829-4820>

References

1. Leucht P, Fischer K, Muhr G, Mueller EJ. Epidemiology of traumatic spine fractures. *Injury*. 2009;40(2):166-172. doi:10.1016/j.injury.2008.06.040.
2. Singh K, Saxena A, Vaccaro AR. Thoracic and lumbar spine trauma. In: *Core Knowledge in Orthopaedics: Spine*. Amsterdam, Netherlands: Elsevier; 2005:290-304. doi:10.1016/B978-0-323-02731-1.50026-6.
3. Schoenfeld AJ, Bono CM. Measuring spine fracture outcomes: Common scales and checklists. *Injury*. 2011;42(3):265-270. doi:10.1016/j.injury.2010.11.040.
4. Stadhouders A, Buckens CFM, Holtslag HR, Öner FC. Are existing outcome instruments suitable for assessment of spinal trauma patients? A review. *J Neurosurg Spine*. 2010;13(5):638-647. doi:10.3171/2010.5.SPINE09128.
5. Öner FC, Jacobs WCH, Lehr AM, et al. Toward the development of a universal outcome instrument for spine trauma: A systematic review and content comparison of outcome measures used in spine trauma research using the ICF as reference. *Spine*. 2016;41(4):358-367. doi:10.1097/BRS.0000000000001207.
6. Öner C, Sadiqi S, Lehr AM, Schroeder GD, Vaccaro AR. The need of validated disease-specific outcome instruments for spine trauma. *J Orthop Trauma*. 2017;31(4):S33-S37. doi:10.1097/BOT.0000000000000945.
7. Sadiqi S, Dvorak MF, Vaccaro AR, et al. Reliability and validity of the English version of the AOSpine PROST (Patient Reported Outcome Spine Trauma). *Spine*. 2020;45(17):E1111-E1118. doi:10.1097/BRS.0000000000003514.
8. Sadiqi S, Lehr AM, Post MW, et al. Development of the AOSpine Patient Reported Outcome Spine Trauma (AOSpine PROST): A universal disease-specific outcome instrument for individuals with traumatic spinal column injury. *Eur Spine J*. 2017;26(5):1550-1557. doi:10.1007/s00586-017-5032-8.
9. Dhakal GR, Sadiqi S, Dhakal R, et al. Reliability and validity of the adapted Nepali version of the AO spine patient reported outcome spine trauma. *J Nepal Health Res Counc*. 2022;19(4):730-739. doi:10.33314/jnhrc.v19i04.3900.
10. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-3191. doi:10.1097/00007632-200012150-00014.
11. Rabin R, de Charro F. EQ-5D: A measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343. doi:10.3109/07853890109002087.
12. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Pol*. 1990;16(3):199-208. doi:10.1016/0168-8510(90)90421-9.

13. Greiner W, Weijnen T, Nieuwenhuizen M, et al. A single European currency for EQ-5D health states. *Eur J Health Econ.* 2003;4(3):222-231. doi:10.1007/s10198-003-0182-5.
14. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil.* 2000;81(12 suppl 2): S15-S20. doi:10.1053/apmr.2000.20619.
15. Altman DG. *Practical Statistics for Medical Research.* Boca Raton, FL: Chapman and Hall/CRC; 1990. doi:10.1201/9780429258589.
16. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428. doi:10.1037/0033-2909.86.2.420.
17. Sadiqi S, Post MW, Hosman AJ, et al. Reliability, validity and responsiveness of the Dutch version of the AOSpine PROST (Patient Reported Outcome Spine Trauma). *Eur Spine J.* 2021; 30(9):2631-2644. doi:10.1007/s00586-020-06554-w.
18. de Leeuw ED. Reducing missing data in surveys: An overview of methods. *Qual Quant.* 2001;35(2):147-160. doi:10.1023/A:1010395805406/METRICS.
19. Hakbijl AJ, van Diemen T, Dompeling S, Oner FC, Post MWM, Sadiqi S. Applicability of the AO Spine Patient Reported Outcome Spine Trauma (PROST) in people with motor-complete spinal cord injury: A cognitive interview study. *Spinal Cord.* 2022;60(10):911-916. doi:10.1038/s41393-022-00829-3.
20. AO Spine outcome instruments. <https://www.aofoundation.org/spine/clinical-library-and-tools/outcome-instruments>. Accessed October 7, 2022.