Less is more

Reliability and measurement error for three versions of the Tampa Scale of Kinesiophobia (TSK-11, TSK-13, and TSK-17) in patients with high-impact chronic pain

Eiger, Bettina; Errebo, Mette; Straszek, Christian Lund; Vaegter, Henrik Bjarke

Published in:
Scandinavian Journal of Pain

DOI:
10.1515/sjpain-2021-0200

Publication date:
2023

Document version:
Final published version

Citation for published version (APA):

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use
This work is brought to you by the University of Southern Denmark. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk

Download date: 15. Sep. 2023
Short Communication

Bettina Eiger*, Mette Errebo, Christian Lund Straszek and Henrik Bjarke Vaegter

Less is more: reliability and measurement error for three versions of the Tampa Scale of Kinesiophobia (TSK-11, TSK-13, and TSK-17) in patients with high-impact chronic pain

https://doi.org/10.1515/sjpain-2021-0200
Received November 1, 2021; accepted February 1, 2022; published online February 22, 2022

Abstract

Objectives: The Tampa Scale of Kinesiophobia (TSK) is a valid and reliable tool to assess somatic focus and activity avoidance in patients. Currently, the test-retest reliability and measurement error for the Danish version is unknown. The aim of the study was to determine standard error of measurement (SEM) and smallest detectable change (SDC) for three Danish lengths of the TSK in patients with chronic pain.

Methods: Waiting-list patients (n = 77) completed the TSK-17 twice from home with a test interval between 7 and 14 days. Based on COSMIN recommendations, the test-retest reliability was estimated using intraclass correlation coefficient (ICC2,1), and measurement error in terms of standard error of measurement (SEMc) and SDC95% were calculated.

Results: All three versions showed good test-retest reliability with ICC2,1-values (CI95%) of 0.86(0.79–0.91), 0.88(0.82–0.92) and 0.87(0.81–0.92) for the TSK-17, TSK-13, and TSK-11. The SEM-values were 3.08, 2.42 and 2.10 respectively and SDC95%-values were 8.53, 6.71 and 5.82.

Conclusions: The Danish versions of TSK-11, TSK-13 and TSK-17 showed good to excellent test-retest reliability. SEM and SDC95% values in patients with chronic pain are reported. The TSK-11 did not show systematic bias between test and retest and may be preferred to minimize responder burden.

Keywords: COSMIN; Danish; kinesiophobia; pain; reliability; Tampa Scale of Kinesiophobia.

Introduction

Chronic pain is a large burden on the individual and society and is the primary reason for disability, with approximately 20% of westerners living with chronic pain [1]. Emotional and cognitive factors have a strong influence on pain and disability [2–5], and fear-avoidance beliefs is one of the factors that has been shown to play a role in the transition from acute to chronic pain [6, 7].

Kinesiophobia was introduced by Kori et al. [8], and defined as “an excessive, irrational, and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or reinjury” and fear of movement has been associated with disability in individuals with chronic pain [4]. Fear of movement can be assessed based on patient reported outcome measures (PROM) with the Tampa Scale of Kinesiophobia (TSK) [9, 10] commonly used.
Although it is debatable what TSK actually measures [11], TSK scores have been shown to predict the degree of disability in patients with chronic pain [12] and the TSK scores correlates with depression, the degree of pain related fear [13–15] and perceived pain intensity [7]. Apart from the original 17-items TSK-17, TSK exists in abbreviated forms, most frequently used is TSK-13 [16–21] and TSK-11 [22–26]. A recently published, modified Delphi study recommends the use of TSK-11 and TSK-17 for identifying fear avoidance beliefs in populations with musculoskeletal pain [27]. TSK is currently available in 14 languages [10, 26, 28–40] and has been translated into Danish [32].

Each translated version represents a new patient reported outcome measure and needs to be psychometrically tested [41, 42]. According to the taxonomy, terminology and definitions from consensus-based standards for the selection of health measurement instruments (COSMIN) test-retest reliability is the extent to which a score of a patient remains unchanged over time at repeated measures, given that the construct of interest remains stable and represents the relative reliability [42]. In addition, when stratifying patients in the clinic or in research, it is important to know the PROMs absolute reliability i.e. measurement error in terms of the standard error of measurement (SEM) and the smallest detectable change (SDC) [41, 43] as expressed at an individual level. Currently the TSK-17 is part of an electronic clinical pain registry (PainData) implemented in specialized interdisciplinary pain centers in Denmark, however, the test-retest reliability and measurement error for the Danish versions (TSK-11, TSK-13 and TSK-17) remains to be investigated in patients with high-impact chronic pain, defined as persistent pain with substantial restriction of life activities lasting 6 months or more [44, 45]. It is essential to investigate if shortened versions are as reliable as the original TSK-17 to decrease the responder-burden. Thus, the aim of this study was to explore the test-retest reliability and measurement error in terms of SEM and SDC for three Danish versions of the TSK [32] in patients with high-impact chronic pain referred to an interdisciplinary pain center.

**Methods and materials**

We conducted the study at the Pain Center, University Hospital Odense, Denmark between November 3rd, 2018, and March 15th, 2019. Through-out this study COSMIN [41–43, 46] guidelines are used.

**Procedure**

Patients with high-impact chronic pain referred to the pain center were invited to participate. As part of the routine clinical data collection approved by the Danish Data Protection Agency (18/35221), all participants answered questions about clinical pain characteristics, adaptations to pain and socio-demographic characteristics through a web-based questionnaire system (PainData) sent via a personal link to the patients’ official inbox. In addition, those who consented to participate in this study were asked to answer the 17 item TSK and average pain intensity one week after the first completion. All questionnaires were answered at home prior to consultations in the pain center.

**Measurements**

**Sociodemographic data**

Sociodemographic data; sex, age, body-mass-index (BMI), marital status and level of education were collected

**The Tampa Scale of Kinesiophobia**

The Tampa Scale of Kinesiophobia (TSK-17) (see Supplementary table 1) [10] consists of 17 items, each rated on a 4-point Likert scale with 1 = “strongly disagree” and 4 = “strongly agree”. Higher scores indicate higher levels of kinesiophobia. The 13 items TSK (TSK-13) version showed that reversed items 4, 8, 12 and 16 had week associations with the total TSK score after factor analysis [16]. The 11 items TSK (TSK-11) showed redundancy for items 9 and 14 after factor analysis [25].

**Pain intensity**

Pain intensity assessed as worst, average, and least in the last 24 h using NRS ranging from 0 (“no pain”) to 10 (“worst imaginable pain”) [47, 48] were assessed at both time points.

**Pain-related disability**

Pain-related disability using the pain disability index (PDI) [49], developed to measure to which degree chronic pain interferes with the ability to engage in various life activities, here used in an abbreviated five question version, with better internal consistency [50]. Each question is answered on an 11-point numerical rating scale from 0–10.
(0 = “no limitations” and 10 = “high degree of limitations due to pain”).

Statistics

All data were analyzed using IBM SPSS Statistics (version 25, IBM Corp., Armonk, NY). Potential differences in pain intensity scores were examined using the Wilcoxon ranks test.

Test-retest reliability

Mean difference with 95% confidence intervals for TSK-scores (test and retest) and intraclass correlation coefficient (ICC_{2,1}) to express the test-retest reliability was calculated. To minimize risk of recall bias retest was determined as minimum 7 days after baseline test and maximum 14 days after to avoid a change in the construct of interest. Excellent reliability was set to >0.90, good 0.75–0.90, moderate 0.5–0.75, and poor <0.5 [51].

Measurement error

Measurement error was estimated by calculating the standard error of measurement (SEM_{agreement}) [42], in which both random and systematic errors are included, by using the formula: \( SEM = SD \times \sqrt{(1 - ICC)} \) where standard deviation \( SD = \sqrt{(SS_{total}/n - 1)} \) [52]; \( SS_{total} \) is sum of squares total, \( n = \) the total number of tests + retest and ICC is the calculated ICC_{2,1}. SEM_{agreement} was converted to SDC95% using the formula: \( SDC95\% = 1.96 \times \sqrt{2} \times SEM \) where SDC95% informs, that it is at an individual level [53, 54].

Results

As illustrated in Figure 1, a total of 213 patients consented to participate in the study approved by the Danish Data Protection Agency (18/35221). However, 105 patients failed to answer the retest altogether or did not have e-Boks, which resulted in 108 included patients. Eighteen patients were excluded due to a time interval >14 days between the two time points, and 13 patients were excluded due to missing items on the TSK-17, resulting in 77 patients included in the analyses.

Population characteristics and descriptive data is shown in Table 1. The average interval between test and retest was 8.4 ± 1.9 days and the pain intensities at the two time points (6.5 ± 1.9 and 6.3 ± 1.7; Wilcoxon rank test: \( p = 0.56 \)) were comparable, indicating stability over time.

TSK scores and test-retest reliability

The scores for the three versions of the TSK at time point one and two are presented in Table 2. The TSK-score at time point 2 (retest) was higher than at time point 1 (test) across all three TSK versions, with differences of 1.0 (TSK-17), 0.8 (TSK-13) and 0.6 (TSK-11), respectively. For the TSK-17 and the TSK-13, the 95% confidence interval of the mean difference, did not include zero within the interval indicating a systematic difference between the test and retest scores. For the TSK-11, the difference between test and retest indicated no systematic difference reflected in the 95% confidence interval of the mean difference, where zero did lie within the interval. As illustrated in Table 2 all three TSK versions showed good to excellent test-retest reliability with ICC_{2,1}-values between 0.86 and 0.88. No floor or Ceiling effect were found in any of the questionnaires as the percentage of maximum or minimum score was 0.00, 1.30 and 1.95 for the TSK-17, TSK-13 and TSK-11 respectively.

Measurement error

As illustrated in Table 2, SEM_{agreement} was 3.08 (TSK-17), 2.42 (TSK-13) and 2.10 (TSK-11) respectively, and the SDC_{ind} was 8.53 (TSK-17), 6.71 (TSK-13) and 5.82 (TSK-11), indicating that on a patient level a change in score of 9 or above on the TSK-17 with a 95% probability is a true change in the kinesiophobia score.

Discussion

In all three versions of TSK we found good to excellent test-retest reliability. However, the mean difference in scores for test and retest indicated systematic bias for the TSK-17 and TSK-13 versions, but not for TSK-11.

Test-retest reliability

The TSK-17 reliability observed in this study is similar to two previous studies that examined the test-retest reliability in patients with chronic pain. In 50 patients, Askary-Ashtiany et al. [55] found an ICC_{2,1}-value of 0.80 over a time interval of 48 h, and in 94 patients, Koho et al. [56] found an ICC_{2,1}-value of 0.9 over a time interval of 24 h. The study by Lamé et al. [57] had an average interval between tests of 52 days, which possibly explains the moderate ICC-value of 0.72 on TSK-17, and 0.74 on TSK-13 [57]. Our result on the TSK-11 is in agreement with the results by
Hapidou et al. [58], whom also tested in a population of 18 patients with chronic pain, with a test-retest interval of 5 days. They found an ICC2,1 value of 0.81 (0.58–0.93). In a group of 264 elderly patients with chronic pain, Larsson et al. [59] found an ICC-value of 0.75, but did not specify the ICC model.

**Measurement error**

For TSK-17, we found an SEM value of 3.08 in our sample of 77 patients with high-impact chronic pain, which is comparable to the SEM of 3.16 found by Woby et al. in a sample of 111 patients with chronic low back pain [25] and the SEM of 3.3 found by Ostelo et al. in a sample of 176 patients with non-specific low back pain with a duration less than 4 months [60]. We found an SDC of 8.53 which is also comparable to the SDC of 9.2 found by Ostelo et al. [60].

For TSK-13, we found an SEM of 2.42 and an SDC of 6.71. To our knowledge SEM and SDC values for the TSK-13 have not previously been reported. Haugen et al. determined a repeatability according to Bland & Altman’s limit of agreement of 8 points in a sample of 52 patients with sciatica due to lumbar disc herniation [35] but did not report SEM or SDC which limits comparison between the studies.

For TSK-11, we found an SEM-value of 2.10. Hapidou et al. reported an SEM-value of 2.41 on the TSK-11 in a small sample of 18 patients [58], which is similar to our finding.

Two studies in samples of patients with chronic low back pain reported an SEM of 2.54 and 1.7, respectively [25, 61]. The short interval between test and retest of 48 h [61] may...
Table 1: Population characteristics (n = 77).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Measurement</th>
<th>Mean (±SD) or number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Women</td>
<td>51 (66.2%)</td>
</tr>
<tr>
<td>Age</td>
<td>Years</td>
<td>49.9 (±14.4), 16–78</td>
</tr>
<tr>
<td>BMI (n = 74)</td>
<td>(Kg/m²)</td>
<td>29.5 (14.7–56.8)</td>
</tr>
<tr>
<td>Marital status</td>
<td>Married/Co-living</td>
<td>57 (74%)</td>
</tr>
<tr>
<td></td>
<td>Single</td>
<td>20 (26%)</td>
</tr>
<tr>
<td>Level of education</td>
<td>Primary school</td>
<td>11 (14.3%)</td>
</tr>
<tr>
<td></td>
<td>Upper secondary school</td>
<td>4 (5.2%)</td>
</tr>
<tr>
<td></td>
<td>Vocational education and training</td>
<td>32 (41.6%)</td>
</tr>
<tr>
<td></td>
<td>Short cycle education</td>
<td>4 (5.2%)</td>
</tr>
<tr>
<td></td>
<td>Vocational bachelor education</td>
<td>15 (19.5%)</td>
</tr>
<tr>
<td></td>
<td>Master’s program</td>
<td>3 (3.9%)</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>7 (9.1%)</td>
</tr>
<tr>
<td>Disability due to</td>
<td>PDI</td>
<td>37 (0–50)</td>
</tr>
<tr>
<td>pain</td>
<td>Average pain</td>
<td>7 (0–10)</td>
</tr>
<tr>
<td></td>
<td>NRSavg (test) (n = 77)</td>
<td>6 (0–10)</td>
</tr>
<tr>
<td>Minimum pain</td>
<td>NRSavg (re-test) (n = 76)</td>
<td>5 (0–10)</td>
</tr>
<tr>
<td>Maximum pain</td>
<td>NRSmin (test) (n = 77)</td>
<td>5 (0–10)</td>
</tr>
<tr>
<td></td>
<td>NRSmax (test) (n = 76)</td>
<td>5 (0–10)</td>
</tr>
<tr>
<td></td>
<td>NRSmax (re-test) (n = 76)</td>
<td>8 (0–10)</td>
</tr>
<tr>
<td>Pain duration</td>
<td>Days: Median (min–max)</td>
<td>2,375 (221–1,477)</td>
</tr>
</tbody>
</table>

Data is described with mean (±SD) for normally distributed data and median (min–max) for non-normally distributed data. BMI, body mass index; PDI, pain disability index; NRS, numeric rating scale; NRSavg, average pain within the latest 24 h; NRSmin, minimum pain the latest 24 h; NRSmax, maximum pain the latest 24 h.

not have limited recall bias leading to a higher ICC and thereby low SEM and SDC values.

Limitations

In total, 105 patients did not fill out the questionnaire a second time and were not included in the study. Maybe it is that they consented to participation, but the burden of filling out the routine clinical data was too much, so they ignored the follow-up questionnaire. Of the eligible 108 patients, 77 patients answered in the given time frame of 7–14 days with-out any missing items, which leaves the sample smaller than the recommended 100 [43, 62, 63]. The excluded patients are slightly older, have a longer history with pain but have similar pain ratings to those included as illustrated in Table 3.

The specialized population of patients with high impact chronic pain selected from a pain clinic limits generalizability. Despite this we choose to investi-gate the reliability of this instrument as it has already been implemented into clinical practice in this population and setting and is part of the PainData registry.

Table 2: Mean (±SD), absolute test-retest difference, ICC2,1, SEMagreement and SDC for TSK-17, TSK-13, and TSK-11.

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD (95%CI)</th>
<th>Mean ± SD (95%CI)</th>
<th>Difference mean ± SD (95%CI)</th>
<th>ICC2,1 (95%CI)</th>
<th>SEMagreement</th>
<th>SDCind</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSK-17</td>
<td>40.1 ± 8.5 (38.2–42.1)</td>
<td>41.2 ± 8.1 (39.3–43.0)</td>
<td>-1.0 ± 4.3 (-2.0 –0.0)</td>
<td>0.86 (0.79–0.91)</td>
<td>3.08</td>
<td>8.53</td>
</tr>
<tr>
<td>TSK-13</td>
<td>31.8 ± 7.3 (30.1–33.4)</td>
<td>32.5 ± 6.8 (31.0–34.1)</td>
<td>-0.8 ± 3.4 (-1.6 –0.3)</td>
<td>0.88 (0.82–0.92)</td>
<td>2.42</td>
<td>6.71</td>
</tr>
<tr>
<td>TSK-11</td>
<td>27.6 ± 6.2 (26.2–29.0)</td>
<td>28.2 ± 5.6 (26.9–29.5)</td>
<td>-0.6 ± 2.9 (-1.3 –0.1)</td>
<td>0.87 (0.81–0.92)</td>
<td>2.10</td>
<td>5.82</td>
</tr>
</tbody>
</table>

TSK, Tampa scale of kinesiophobia; SD, standard deviation; 95% CI, confidence interval 95%; ICC2,1, intra class correlation – two-way random effects, total agreement; SEMagreement standard error of measurement with total agreement; SDCind, smallest detectable change = SDC95%.
Similar to previous studies [33, 59], we chose to exclude patients with missing items, which could affect results. A total of 13 (14.4%) patients were excluded due to missing items, which is close to the limit of maximum 15% recommended by De Vet et al. [64]. It corresponds to the results of Larsson et al. [59] (14.2%) but is lower than Lundberg et al. [33] (18.5%) and Koho et al. [56] (29.9%).

Implications

After treatment it is important to know if the TSK can identify a change in kinesiophobia. SDC informs if the change is over and above measurement error and is in the same unit as the questionnaire.

Previous studies have recommended TSK-11 when assessing the degree of kinesiophobia in patients with chronic pain, since it has been found valid and reliable [23, 25, 59]. The shorter version is less time consuming for both patient and health care professional. We found only [23, 25, 59].

After treatment it is important to know if the TSK can identify a change in kinesiophobia. SDC informs if the change is over and above measurement error and is in the same unit as the questionnaire.

Research funding: Authors state no funding involved.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Conflict of Interest: There are no actual or potential conflicts of interest for any of the authors.

Informed Consent: Written informed consent (electronic) was obtained from all patients included in this study.

Ethical Approval: The Danish Data Protection Agency approved the data collection (ref. no. 18/35221), and the conduct of this study complied with the Declaration of Helsinki. No ethics approval was needed (Act on Research Ethics Review of Health Research Projects, October 2013, Section 14.2).

References

36. De Souza FS, Marinho Cda S, Siqueira FB, Maher CG, Costa LOP. Psychometric testing confirms that the Brazilian–Portuguese adaptations, the original versions of the fear-avoidance beliefs questionnaire, and the Tampa scale of kinesiophobia have similar measurement properties. Spine 2008;33:1028–33.


Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/npjpain-2021-0200).