Psychometric properties of the Tampa Scale of Kinesiophobia (TSK-11) among older people with chronic pain

Caroline Larsson, RPT, MSc1, Eva Ekvall Hansson, RPT, PhD2, Kristina Sundquist, MD, PhD1,3, and Ulf Jakobsson, RN, PhD1

1Center for Primary Health Care Research, Lund University/Region Skåne, Malmö, Sweden, 2Department of Clinical Sciences in Malmö/Family Medicine, Lund University, Malmö, Sweden, and 3Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA

Abstract

Objectives: The study aimed to test the construct validity, factor structure and reliability of the 11-item version of the Tampa Scale for Kinesiophobia (TSK-11, Swedish version) among older people (65+) with chronic pain. Design: Methodological study. Subjects: 433 participants with chronic pain (mean age 74.8, 65–98 years) completed postal questionnaires. 264 of the participants completed a test–retest assessment. Methods: Construct validity was evaluated through corrected item-total correlations. Convergent validity was analyzed by correlations with activity/activities of daily living (ADL) dependence, pain intensity and physical activity (all of which are constructs related to kinesiophobia according to fear-avoidance theories). Factor structure was tested through confirmatory factor analysis. Reliability was assessed with Cronbach’s α and test–retest reliability, analyzed by intra-class correlation coefficient (ICC) and weighted κ coefficient analysis. Results: Evidence of convergent validity was shown by significant positive correlations with activity/ADL dependence (r = 0.20) and pain intensity (r = 0.31), and a significant negative correlation with physical activity (r = −0.38). Confirmatory factor analysis showed that both one- and two-factor-solutions were possible. Cronbach’s α coefficients ranged between 0.74 and 0.87. Test–retest analysis showed strong agreement regarding ICC (r = 0.75, 95% confidence interval 0.64–0.82). The weighted κ coefficients for the individual items showed fair to moderate reliability. Conclusion: The Swedish version of TSK-11 had acceptable construct validity, factor structure, and reliability and, hence, can be considered suitable for older people with chronic pain.

Introduction

Kinesiophobia, defined as “an excessive, irrational, and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or reinjury” (Kori, Miller, and Todd, 1990), is an important factor in the development of chronic pain and a strong predictor of disability (Turk and Wilson, 2010; Vlaeyen and Linton, 2000).

According to the framework of fear-avoidance theories, a painful experience triggered during movement, when interpreted as being threatening, can lead to kinesiophobia. Avoidance behavior is a consequence of kinesiophobia and is seen as a natural response to injury but has, if prolonged, negative effects on both physical and psychological functions. Daily activities and functional capacity may be reduced to avoid pain, leading to decreased physical activity, disuse, disability and further chronicity of pain (Vlaeyen and Linton, 2000). Kinesiophobia is therefore an important domain to consider when assessing psychosocial characteristics of pain.

An instrument commonly used for measuring kinesiophobia is the Tampa Scale for Kinesiophobia (TSK-17) (Kori, Miller, and Todd, 1990). The TSK-17 consists of 17 items requesting patients to rate their responses to 17 different statements relating to fear of movement and (re)injury. It is widely used among patients with a variety of types of pain (Bunketorp, Carlsson, Kowalski, and Stener-Victorin, 2005; Lundberg, Styf, and Jansson, 2009; Mintken et al, 2010; Swinkels-Meewisse et al, 2003; Woby, Roach, Urmston, and Watson, 2005), and has been translated into several languages (Damsgard, Fors, Anke, and Roe, 2007; Gomez-Perez, Lopez-Martinez, and Ruiz-Parraga, 2011; Vlaeyen et al, 1995), including Swedish (Lundberg, Styf, and Carlsson, 2004).

Studies have evaluated the psychometric properties of the TSK-17 among patients with different pain diagnoses, and have shown acceptable content and face validity (Lundberg, Styf, and Carlsson, 2004), internal consistency (Cronbach’s α: 0.68–0.91) (Lundberg, Styf, and Carlsson, 2004; Vlaeyen et al, 1995) and test–retest reliability (ICC: 0.64–0.91) (Lamé et al, 2008; Lundberg, Styf, and Carlsson, 2004; Swinkels-Meewisse et al, 2003).

Theoretic rationales for associations between kinesiophobia and pain intensity, disability and physical activity have been hypothesized in fear-avoidance models (Vlaeyen and Linton, 2000) and confirmed in subsequent studies. These studies indicate positive correlations between TSK and disability (r = 0.30–0.51) and pain intensity (r = 0.23–0.27), and a negative correlation between TSK and physical activity (r = −0.48) (Boersma and Linton, 2005; Clark, Kori, and Brockel, 1996; Elfving, Andersson, and Grooten, 2007; Koho et al, 2011; Turk and Wilson, 2010; Vlaeyen and Linton, 2000; Woby, Roach, Urmston, 2005).
and Watson, 2005). However, the construct validity and factor structure of the TSK-17 have been questioned. Consequently, different factor structures and shorter versions of the instrument have been proposed. An example of a shortened version that has received increasing attention due to its improvements in user-friendlyness, administration and scoring time is a 11-item version named TSK-11 (Roelofs et al, 2007; Woby, Roach, Urmston, and Watson, 2005).

Since the prevalence of chronic pain is high among older people and many older people live with untreated or only partially treated pain (Gibson and Lussier, 2012; Jakobsson, 2010), studies about kinesiophobia is especially important in this group. However, such studies are very sparse and thus highly needed. When using health measurement scales it is important that the assessment is carried out with instruments that are psychometrically sound for the specific group of patients (e.g. in terms of language and age). Previous studies of the psychometric properties of the Tampa Scale for Kinesiophobia have only been performed among middle-aged people. It seems that no study has evaluated its psychometric properties among older age groups. Thus, further testing is needed before the Tampa Scale for Kinesiophobia can be used among older populations and the specific contexts and the heterogeneity of groups of ‘‘older people’’ must be considered. Furthermore, age-related changes (e.g. visual impairments and changes in cognitive processing) may impair older persons’ comprehension of long assessments (Malmstrom and Tait, 2010). Thus, the TSK-11 may be preferable to the 17-item version when used among older people.

The objective of the present study was to psychometrically test the construct validity, factor structure and reliability of the TSK-11(Swedish version) among older people (age 65+ years) with chronic pain.

**Methods**

**Subjects**

This study was carried out in 2011 and included people aged 65 years and older. The participants were randomly selected using a Swedish national register of inhabitants (SPAR) by a government-engaged company (Infodata). The sampling frame was Sweden’s whole population in the selected age group. A total of 2000 questionnaires were mailed out and 1141 were sent back in usable condition (filled in) by respondents aged 65–103 years. Twelve questionnaires were returned without having been filled in for the reason ‘‘address unknown’’ and 13 people were reported to be dead. The response rate was 57.8%. Of the 1141 respondents, 433 (37.9%) reported suffering from pain for more than 3 months and they comprised the sample used for analyses in the present study. The only variable available for analysis of the non-responders was gender, for which no significant difference ($p = 0.322$) was found between participants (53.2% women) and non-participants (55.8% women).

**Procedure and measurements**

Data were collected by questionnaires that were distributed by post together with an accompanying letter explaining the aim and procedure of the project. The questionnaires were requested to be sent back in enclosed self-addressed, prepaid envelopes.

A test–retest questionnaire containing the TSK-11 was sent to a subsample of the respondents reporting chronic pain ($n = 336$) two weeks following their response to the baseline questionnaire. Two hundred and sixty-two of the 336 participants filled in and returned the follow-up questionnaire, resulting in a response rate of 78% for the test–retest analyses.

The baseline questionnaire contained questions about gender, age, housing and living conditions, and presence of chronic pain.

**Pain**

Pain was assessed using 3 items extracted from the brief screening version of the Multidimensional Pain Inventory (Swedish version), which contains questions about pain intensity, duration and location (Bergström, Jensen, Linton, and Nygren, 1999).

Pain intensity was measured using the item ‘‘Rate the average level of your pain during the last week’’ responding to a 6-point Likert scale with answers ranging from ‘‘No pain at all’’ (value of 1 point) to ‘‘Tremendous amount of pain’’ (value of 6 points). Duration was measured in years with pain. Chronic pain was defined as having pain in 3 months.

**Kinesiophobia**

In this study, we used a shortened version of the Tampa Scale for Kinesiophobia, based on an 11-item version (TSK-11) proposed by Woby, Roach, Urmston, and Watson (2005). The 17-item version of The Tampa scale of Kinesiophobia has previously been translated into Swedish (through back-forward translation) (Lundberg, Styf, and Carlsson, 2004) and the version used in this study is an abbreviation of the existing Swedish version, excluding 6 of the original items in the same way as in TSK-11. The TSK-11 excludes items 4, 8, 9, 12, 14, and 16 from the original 17-item instrument. Like the 17-item version, TSK-11 is anchored with the answers ‘‘strongly disagree’’, which scores 1 point, and ‘‘strongly agree’’, which scores 4 points. The total summary score is calculated and can range between 11 and 44 points. A high score indicates strong fear of movement/reinjury (i.e. high kinesiophobia).

A two factor solution has been used in this study. The subscale Somatic Focus (TSK-SF; belief in underlying and serious medical problems) comprises items 3, 11, 6, 7, 5 and has a total score that ranges from 5 to 20 points. The second subscale Activity Avoidance (TSK-AA; belief that activity may result in (re)injury or increased pain) comprises items 1, 2, 10, 13, 15, and 17 and has a total score of 6–24 points (Clark, Kori, and Brockel, 1996; Roelofs et al, 2007).

The TSK-11 has been found to have good construct validity (correlation ($r$) of the TSK-11 with disability = 0.51 and with pain intensity = 0.27) and internal consistency (Cronbach $\alpha = 0.77–0.91$) when tested across various nationalities and pain populations (Roelofs et al, 2007; Woby, Roach, Urmston, and Watson, 2005). However, all previous tests were performed among young and middle-aged subjects (age 20–65 years).

**Physical activity**

Physical activity has been defined as bodily movement produced by skeletal muscles that requires energy expenditure and that can be categorized into occupational, sports, conditioning, household or other activities (Caspersen, Powell, and Christenson, 1985). Physical activity was measured with Grimby’s activity scale, developed to evaluate self-rated physical activity among older people (Saltin and Grimby, 1968). Levels of physical activity were classified using the question ‘‘How physically active do you think you have been during the last six months?’’ The alternative responses were: 1 = ‘‘Hardly any physical activity’’, 2 = ‘‘Mostly sedentary’’, 3 = ‘‘Lighter physical exercise, around 2–4/h/week’’, 4 = ‘‘More strenuous exercise, 1–2 times/week’’, 5 = ‘‘More strenuous exercise, at least 3/h/week’’ and 6 = ‘‘Hard regular training, several times/week’’.
The physical activity scale has been psychometrically evaluated in older people and has demonstrated good construct validity when validated against measures of physical performance (Aires, Selmer, and Thelle, 2003; Frändin and Grimby, 1994; Grimby, 1986).

Disability
Disability is an umbrella term including limitations in activities of daily living (ADL) as well as impairments and participation restrictions (World Health Organization, 2001). In this study, we used disability synonymous to activity dependence, which is classified as a limit in ADL according to the International Classification of Functioning (ICF). To measure disability as activity dependence we used the ADL staircase which is a Swedish instrument developed from the Katz ADL Index (Hultcr Asberg, 1990; Katz and Akpom, 1976). The instrument describes the degree of dependence/independence in 10 activities of ADL. The activities are divided into the dimensions: Personal ADL (bathing, dressing, toilet use, transferring, continence, and feeding); and Instrumental ADL (cooking, transportation, grocery shopping, and cleaning). ADL ability can be assessed by self-assessment, observation or testing, and all activities are graded as independent or dependent. Dependence/independence is then graded using a 10-point scale, where 10 represents total dependence and 0 represents independence. ADL staircase has previously been used to measure disability among older people with chronic pain (Widar and Ahlström, 2002). When psychometrically tested on older people, the scale demonstrated acceptable construct validity and high values for reliability (Cronbach’s α for internal consistency; 0.85–0.90) (Jakobsson, 2008).

Ethical considerations
The study was conducted in accordance with the basic ethical principles of medical research (MRC, 2000) and was approved by the Regional Ethical Review Board in Lund (approval number 2010/683).

Statistical analyses
For demographic data, mean, standard deviation, range and percentiles were used as descriptive measures. The TSK-11 was evaluated for construct validity and reliability. Floor and ceiling effects were assessed and were considered to be present if more than 15% of the sample chose the highest or lowest option, respectively. Item response rate and overall response rate were also calculated to assess the response frequency of the instrument.

Factor structure
The factor structure of the instrument was examined by confirmatory factor analysis using LISREL statistical software version 8.80. Based on previous research on factor structures of the TSK-11, a two-factor solution (TSK-AA, TSK-SF) and a one-factor solution (TSK-11) were tested (Clark, Kori, and Brockel, 1996; Roelofs et al, 2007). Confirmatory factor analysis was used to statistically determine whether previously established models fitted the shared variance of the measured variables using Spearman’s rank order correlation matrix. To measure the model’s fit, several goodness-of-fit tests were performed. We calculated the goodness of fit index (GFI), adjusted goodness of fit index (AGFI), normed fit index (NFI), and comparative fit index (CFI), all of which can range between 0 (no fit at all) and 1 (perfect fit). They have no absolute threshold levels for acceptability, but values of 0.85–0.95 are generally considered to indicate acceptable fit to the model and values of 0.90–0.97 good fit. We also calculated the root mean square error of approximation (RMSEA), which is considered acceptable when its value is ≤0.08; however, a value of ≤0.05 is desirable (Schermelleh-Engel, Moosbrugger, and Müller, 2003).

Validity
Validity was evaluated through analysis of construct validity. Convergent validity (how closely the scale is related to other variables/measures of the same construct) (Streiner and Norman, 2008) was analyzed using Spearman’s rank order correlation. Our a priori hypotheses were based on the theoretical framework presented in the cognitive behavioural fear avoidance model by Vlaeyen and Linton (2000). We hypothesized a positive correlation between kinesiophobia (TSK) activity/ADL dependence (ADL staircase), and pain intensity (numerical rating scale). In addition a negative correlation was hypothesized between kinesiophobia (TSK) and physical activity (Grimby’s activity scale) (Boersma and Linton, 2005; Clark, Kori, and Brockel, 1996; Elving, Andersson, and Grooten, 2007; Kho et al, 2011; Turk and Wilson, 2010; Vlaeyen and Linton, 2000; Woby, Roach, Urmston, and Watson, 2005).

Corrected item-total correlations reflect the construct validity and homogeneity of an instrument and were used to measure how well one individual item correlated with the total score for the remaining items when it is omitted. Item-total correlation analysis was performed to check whether any item in the set of tests was inconsistent with the averaged behavior of the others. It is recommended that a value of <0.2 be interpreted as an indication that the individual item is not correlating well with the other items and should therefore be discarded from the instrument (Streiner and Norman, 2008). A value >0.2 is desirable and values >0.3 can be considered more suitable.

Reliability
The reliability of the TSK was tested with regard to internal consistency and test–retest reliability. Cronbach’s α was used to measure internal consistency (Cronbach, 1951). Generally, values of 0.7–0.9 are considered to indicate good reliability. Values <0.7 could indicate that some of the items are too heterogeneous. On the other hand, a value >0.9 may indicate a high level of item redundancy (Schermelleh-Engel, Moosbrugger, and Müller, 2003).

Test–retest reliability (i.e. correlation agreement between total TSK values at the two assessment time points – baseline and follow-up at 2 weeks) was analyzed by determining the ICC and CIs for the total score of the instrument, where a value of 1 represents perfect agreement and 0 no agreement at all (Streiner and Norman, 2008). ICC values of 0.2–0.4 indicate poor agreement; values of 0.3–0.4 indicate fair agreement; values of 0.5–0.6 indicate moderate agreement; values of 0.7–0.8 indicate strong agreement; and values >0.8 indicate almost perfect agreement. Agreement for the individual items in the test–retest analysis was assessed by weighted κ analysis (Streiner and Norman, 2008). κ values of ≤0.20 represent poor reliability; values of 0.21–0.40 indicate fair reliability; values of 0.41–0.60 indicate moderate reliability; values of 0.61–0.80 represent good reliability; and values >0.80 are considered to represent very good reliability (Altman, 1990). Data were analyzed using PASW 20.0, Vassar Stats (http://www.vassarstats.net) and LISREL statistical software version 8.80.

Results
The demographic characteristics of the respondents are presented in Table 1. Of the total sample at baseline 38.5% (n = 433) reported having had pain for more than 3 months, with a mean
duration of 10.2 years (SD 12.2). The respondents with chronic pain were aged 65–98 years, with a mean age of 74.8 years, and 63.5% (275) were women.

For mean total TSK-11 score, differences were found between those who were living in special accommodation (mean = 33.8 (SD 6.99) and those living in the community (mean = 22.3 (SD 8.21)) (p = 0.001). There was no significant difference in mean total TSK-11 score between men (mean = 22.66 (SD 8.22)) and women (mean = 22.4 (SD 8.46)) (p = 0.771) (data not shown in tables). Item response rates ranged between 82.9 and 87.1% (Table 2), and a total score for the instrument could be calculated for 75.8% of the respondents. Of the respondents 5.3% received the highest score of 44 points and 0.7% received the lowest score of 11 points. Thus no floor or ceiling effects were found for the total score.

Factor structure

The results of the confirmatory factor analysis are presented in Table 3. For the TSK-11, the factor loadings ranged between 0.57 and 0.83 for the two-factor solution and between 0.49 and 0.82 for the one-factor solution. The factor loadings were slightly higher for the two-factor solution (C21 x2 = 255.8; df, 43), compared to the one-factor solution (C21 x2 = 339.9; df, 44). A two-factor model provides a significantly better fit to the data than a one-factor model, as the difference in chi-square values between the one-factor and two-factor models (about 84) is itself chi-square distributed, with the number of degrees of freedom being the difference between the df values for the two models (i.e. 1). Furthermore, the goodness of fit analysis also showed a slightly better fit for the two-factor solution (Table 4).

Validity

Corrected item total correlations were 0.34–0.74 for the TSK-11 and 0.36–0.55 and 0.57–0.72 for the TSK-SF and TSK-AA subscales, respectively (Table 2). Results for convergent validity are presented in Table 5. The total score for the TSK-11 was significantly and positively correlated with pain intensity (single item) and the ADL staircase, but was negatively correlated with Grimby’s activity scale. The TSK-SF and TSK-AA subscales showed similar but slightly weaker correlations with the ADL staircase and Grimby’s activity Scale. Grimby’s activity scale was
confirmatory factor analysis. Evidence of convergent validity was shown by significant correlations with external measures. This study aimed to evaluate psychometric properties of an 11-item Swedish version of the Tampa Scale of Kinesiophobia (TSK-11). The internal consistency of the TSK-11 was found to be acceptable for all factors (Table 2). Cronbach’s α coefficients for total score, TSK-SF and TSK-AA were 0.87, 0.74, and 0.85, respectively (data not shown in tables). Cronbach’s α coefficients if an item was deleted ranged between 0.67 and 0.70 for TSK-SF and between 0.81 and 0.84 for TSK-AA. Test–retest analysis of total TSK score showed strong agreement regarding the ICC (r = 0.747 [95% CI, 0.64–0.82]). Agreement for the individual items, assessed with weighted κ coefficients, showed moderate reliability (range 0.41–0.58), with the exception of items 1 and 6, which showed fair reliability (Table 2).

Table 3. Factor loadings of the one-factor and two-factor models of the TSK-11, as obtained by confirmatory factor analysis.

<table>
<thead>
<tr>
<th>Item</th>
<th>Total TSK-11</th>
<th>Somatic focus (SF)</th>
<th>Activity avoidance (AA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’m afraid that I might injure myself if I exercise</td>
<td>0.645</td>
<td>0.657</td>
<td></td>
</tr>
<tr>
<td>If I were to overcome it, my pain would increase</td>
<td>0.642</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My body is telling me I have something dangerously wrong</td>
<td>0.491</td>
<td>0.582</td>
<td></td>
</tr>
<tr>
<td>People aren’t taking my medical condition seriously enough</td>
<td>0.522</td>
<td>0.602</td>
<td></td>
</tr>
<tr>
<td>My accident has put my body at risk for the rest of my life</td>
<td>0.624</td>
<td>0.718</td>
<td></td>
</tr>
<tr>
<td>Pain always means I have injured my body</td>
<td>0.597</td>
<td>0.634</td>
<td></td>
</tr>
<tr>
<td>Simply being careful that I do not make any unnecessary movement is</td>
<td>0.718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the safest thing I can do</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I wouldn’t have this much pain if there wasn’t something potentially</td>
<td>0.517</td>
<td>0.571</td>
<td></td>
</tr>
<tr>
<td>dangerous going on in my body</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain lets me know when to stop exercising so that I don’t injure my</td>
<td>0.665</td>
<td>0.694</td>
<td></td>
</tr>
<tr>
<td>self</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can’t do all the things normal people do because it’s too easy for</td>
<td>0.816</td>
<td>0.833</td>
<td></td>
</tr>
<tr>
<td>me to get injured</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No one should have to exercise when he/she is in pain</td>
<td>0.607</td>
<td>0.626</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Fit of the data according to confirmatory factor analyses.

<table>
<thead>
<tr>
<th>Item</th>
<th>Total TSK-11</th>
<th>Somatic focus (SF) + activity avoidance (AA)</th>
<th>Criteriona</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodness of fit index (GFI)</td>
<td>0.877</td>
<td>0.909</td>
<td>≥0.95</td>
</tr>
<tr>
<td>Adjusted goodness of fit index (AGFI)</td>
<td>0.816</td>
<td>0.859</td>
<td>≥0.90</td>
</tr>
<tr>
<td>Normed fit index (NFI)</td>
<td>0.913</td>
<td>0.934</td>
<td>≥0.95</td>
</tr>
<tr>
<td>Comparative fit index (CFI)</td>
<td>0.923</td>
<td>0.945</td>
<td>≥0.97</td>
</tr>
<tr>
<td>Root mean square error of approximation (RMSEA)</td>
<td>0.124</td>
<td>0.103</td>
<td>≤0.05</td>
</tr>
</tbody>
</table>

aSchermelleh-Engel, Moosbrugger, and Müller (2003).

Table 5. Correlation matrix for TSK-11, pain intensity, ADL staircase and Grimby’s activity scale.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>TSK-11</th>
<th>TSK-SF</th>
<th>TSK-AA</th>
<th>Pain intensity</th>
<th>ADL staircase</th>
<th>Grimby’s activity scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSK-11</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSK-SF</td>
<td>0.858</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSK-AA</td>
<td>0.937</td>
<td>0.641</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain intensitya</td>
<td>0.306</td>
<td>0.342</td>
<td>0.248</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADL Staircase</td>
<td>0.197</td>
<td>0.155</td>
<td>0.177</td>
<td>0.166</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Grimby’s activity scale</td>
<td>−0.378</td>
<td>−0.307</td>
<td>−0.349</td>
<td>−0.188</td>
<td>−0.406</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Significant p values (<0.05) are presented in bold.

aPain intensity = “average level of pain in the last week”, measured using a 5-point Likert scale with answers ranging from “No pain at all” to “Tremendous amount of pain”.

significant correlated with both TSK-11 and TSK-AA. Pain intensity showed the strongest correlation with TSK-SF.

Reliability

The internal consistency of the TSK-11 was found to be acceptable for all factors (Table 2). Cronbach’s α coefficients for total score, TSK-SF and TSK-AA were 0.87, 0.74, and 0.85, respectively (data not shown in tables). Cronbach’s α coefficients if an item was deleted ranged between 0.67 and 0.70 for TSK-SF and between 0.81 and 0.84 for TSK-AA. Test–retest analysis of total TSK score showed strong agreement regarding the ICC (r = 0.747 [95% CI, 0.64–0.82]). Agreement for the individual items, assessed with weighted κ coefficients, showed moderate reliability (range 0.41–0.58), with the exception of items 1 and 6, which showed fair reliability (Table 2).

Discussion

This study aimed to evaluate psychometric properties of an 11-item Swedish version of the Tampa scale of kinesiophobia. Evidence of convergent validity was shown by significant correlations with external measures. Confirmatory factor analysis showed that both one- and two-factor solutions were possible. Reliability was supported by acceptable values for Cronbach’s α and strong agreement regarding ICC. Weighted κ coefficients for the individual items showed fair to moderate reliability.

Factor structure

Both 1 and 2 factor solutions were tested, and the 2-factor solution was found to have slightly better goodness-of-fit. Previous studies have used factor structures ranging from one to five (Lundberg, Grimby-Ekman, Verbunt, and Simmonds, 2011). This variation may have been caused of an unstable fracture structure, but may also be due to differences between populations or differences in diagnosis.

For TSK-11, it was previously implied that the kinesiophobia construct consists of two underlying constructs (namely somatic focus and activity avoidance) (Clark, Kori and Brockel, 1996; Roelofs et al, 2007; Woby, Roach, Urmston, and Watson, 2005). This is supported by the results of this study, which demonstrate that a two-factor TSK-11 solution also provides a good fit among older people with chronic pain. In the present study both subscales correlated strongly with the total score, but only moderately with each other. This was shown previously...
(French et al, 2007; Tkachuk and Harris, 2012) and indicates that both subscales without doubt belong to the overlying kinesiophobia construct, but are also two clear underlying constructs. However, although there are small statistical differences between the models (e.g. in goodness-of-fit tests and the magnitudes of the factor loadings) in favor of the two-factor solution, they are in practice largely equivalent. Thus, both versions may be usable, depending on whether the interest of the study lies in the underlying aspects of kinesiophobia (Somatic Focus and Activity Avoidance) or in assessing general levels of kinesiophobia.

Validity
Acceptable results were obtained for item response rate and floor/ceiling effects. Evidence for construct validity measured as convergent validity and evaluated with respect to external measures.

Convergent validity was supported by positive associations between the TSK scales (TSK-11, TSK-AA, and TSK-SF) and the ADL staircase. Although the associations in our study can be considered weak ($r = 0.16$–$0.20$), an extensive amount of evidence supports a strong association between Kinesiophobia and disability (Bergbom, Boersma, and Linton, 2012; Leeuw et al, 2007; Vlaeyen and Linton, 2000). For example, prospective studies have shown kinesiophobia to be a significant predictor of disability (Kamper et al, 2012; Nieto, Miró, and Huguet, 2009). In line with our results more moderate correlations between kinesiophobia (TSK-17) and disability ($r = 0.30$) has also been reported (i.e. among patients with chronic back pain) (Luning Bergsten, Lundberg, Lindberg, and Elfving, 2012). For the shortened 11-item version (TSK-11) both strong ($r = 0.51$) and weak correlations ($r = 0.24$) have been found (Gomez-Perez, Lopez-Martinez, and Ruiz-Parraga, 2011; Woby, Roach, Urmston, and Watson, 2005). However, the weak correlations in our result for the ADL-staircase make it difficult to support previous findings of others suggesting a stronger relationship between kinesiophobia and disability. Our findings suggest that though there is a relationship, these two concepts are not similar in this sample. It may indicate that there is little support for convergent validity regarding correlations to the ADL-Staircase. One explanation for the variation in the results is that disability is an umbrella concept (which includes ADL dependence). Moreover the disability construct is poorly defined in fear-avoidance theories and hence a large variety of disability measures have been used, making it difficult to compare results between studies. Furthermore, the weak correlation in our result can also be explained by a theoretically linear relationship for the lower values of ADL dependence and TSK-11. However, there is most likely not a linear relationship for the upper parts of the scale. This is due to when the person gets more and more disabled, the more immobilized the person also gets, and lower levels of kinesiophobia can thus be assumed.

Pain intensity was significantly, but moderately correlated ($r = 0.25$–$0.34$) with reductions in the TSK scores (TSK-11, TSK-AA, and TSK-SF). In accordance with previous research ($r = 0.24$) (Roelofs et al, 2007), the TSK-SF ($r = 0.34$) showed the strongest correlation with pain intensity of the three subscales in the analysis of convergent validity. That pain intensity is important in the development of disability has previously been shown among younger samples (Leeuw et al, 2007). Correlations of kinesiophobia with pain intensity in the present study suggest that this also is the case among older people with chronic pain.

As hypothesized, physical activity was found to be moderately associated with kinesiophobia ($r = 0.31$–$0.38$). In agreement with previous studies ($r = 0.23$–$0.48$) showing that low levels of physical activity are associated with higher levels of kinesiophobia (Elfving, Andersson and Grooten, 2007; Koho et al, 2011) a negative correlation was found also in our results. Combining the tests of validity performed in this study, the combined results support the validity of TSK-11 when used among older people with chronic pain.

Reliability
The reliability of the TSK-11 can be considered acceptable regarding both internal consistency and test–retest reliability. Also, the Cronbach’s α coefficients for the total score and when items were deleted were all within the acceptable range, although the TSK-SF values were low. A recent review evaluating different versions (including both TSK-17 and TSK-11) among a variety of subjects presented α values for internal consistency for the total score of $0.70$–$0.86$ were the α values of most studies were found in the lower end of the range (Lundberg, Grimby-Ekman, Verbunt, and Simmonds, 2011). In this study, the values of Cronbach α were slightly higher ($0.74$–$0.87$), indicating that the 17-item version is more heterogenic and that the 11-item version may therefore be a better option for measuring kinesiophobia, due to its advantages of brevity. However, the difference in α values between the results from this study and previous results is small and cannot be used as criteria for which version to choose.

It must be noted that the inter-item correlation analysis yielded acceptable but relatively low κ values, with the values for two items (1 and 6) being slightly lower than those for the other items. All other items presented with values in the moderate range (0.41–0.58) (Table 2). Previous studies of TSK-11 have used shorter time intervals between baseline and follow-up assessments (e.g. $74 \pm 14$ h) (Woby, Roach, Urmston, and Watson, 2005) compared to the present study (2 weeks). The length in the time interval could therefore partly explain the relatively low κ values. However, the ICC value in this study ($r = 0.75$) indicates strong agreement between the two assessments, in line with ICC values previously presented in psychometric testing of both the TSK-17 and TSK-11 (Lamé et al, 2008; Lundberg, Grimby-Ekman, Verbunt, and Simmonds, 2011; Lundberg, Styf, and Carlsson 2004; Woby, Roach, Urmston, and Watson, 2005). Thus, the reliability of the TSK-11 can be considered acceptable for its use among older people with chronic pain.

Study limitations
In this study the response rate can be considered relatively low (57%), this needs to be considered when interpreting the result. A low response rate, if the drop out is systematic, is a threat to the validity of a study. However the response rate of this study may therefore be a better option for measuring kinesiophobia, due to its advantages of brevity. Another limitation is that TSK was initially developed to measure neuropathic) were analyzed separate. However, this study gives a broad view of older people and recent studies have supported the reliability and validity of TSK-11 for heterogeneous chronic pain populations across medical diagnoses (Roelofs et al, 2007).

Conclusion
This study supports the Swedish version of TSK-11, as a valid and reliable measure of kinesiophobia to use among older people with
chronic pain. The TSK-11 had acceptable construct validity, factor structure and test–retest reliability. Factor analysis showed that both a one- and two-factor solutions may be useful to researchers and clinicians, depending on whether the interest lies in the underlying aspects (Somatic Focus and Activity Avoidance) or in general levels of kinesiophobia. This short version is offering the advantages of brevity, reducing administration time and patient burden.

Acknowledgements

The authors wish to acknowledge the respondents for participating in the study. We are also most grateful to Stephen Gilliver for revising the English in this manuscript and to Per Condelius for help with the data collection. This study was funded through generous grants from the King Gustav V. and Queen Victoria's Foundation of Freemasons, the Gyllenstiernska Krapperup Foundation, and Ragnar and Einar Lundström’s Foundation.

Declaration of interest

The authors report no conflicts of interest.

References


Eflving B, Andersson T, Grooten WJ 2007 Low levels of physical activity in back pain patients are associated with high levels of fear-avoidance beliefs and pain catastrophizing. Physiotherapy Research International 12: 23–74.


Hulten Åsberg K 1990 ADL-trappan, 1st edn. Lund, Studentlitteratur.


