Agreement between measures of “walking-related knee pain” in knee osteoarthritis: a cross-sectional study

L Klokker1*, R Christensen1,2, R Osborne3, J Aaboe3, H Bliddal1, M Henriksen1

Abstract

Introduction

Modest effect sizes of clinical pain intervention studies have caused a need for evaluation of the applied methodology, including identification of treatment response indicators. Comparing two measurement instruments helps to identify the underlying constructs, which is important in interpreting the results appropriately, in research as well as in practice. This study aimed at assessing the agreement between a performance measure (walking task including pain assessment) and a patient reported outcome (the Knee Injury and Osteoarthritis Outcome Score), in a population of patients with knee osteoarthritis.

Materials and methods

Cross-sectional data from 143 patients with knee osteoarthritis included in a prospective weight loss study were analysed. All participants completed the Knee injury and Osteoarthritis Outcome Score within one week prior to rating their target knee pain on a 100 mm visual analogue scale after walking 150–200 m at a self-selected pace in a gait laboratory. The Knee injury and Osteoarthritis Outcome Score pain subscale and item 5 of the Knee injury and Osteoarthritis Outcome Score pain subscale (“amount of knee pain experienced during walking on a flat surface in the last week”) were selected for analysis. Distributions of visual analogue scale scores within the Knee injury and Osteoarthritis Outcome Score pain subscale 05 response categories were described and compared using Spearman correlation. Agreement was estimated using Limits of Agreement.

Results

There was a moderate correlation between visual analogue scale and Knee injury and Osteoarthritis Outcome Score pain subscale 05 (r = 0.5, p < 0.001), with a wide range of visual analogue scale scores within the Knee injury and Osteoarthritis Outcome Score response categories. Generally, higher pain scores were reported with the Knee injury and Osteoarthritis Outcome Score pain subscale than with the performance test, with a mean difference of 18.8 (SD 16.6), and Limits of Agreement from –13.6 to 51.3.

Conclusion

Disagreement between the performance measure and Knee injury and Osteoarthritis Outcome Score pain subscale, together with the moderate correlation of visual analogue scale and Knee injury and Osteoarthritis Outcome Score pain subscale 05 scores, suggests a difference in the underlying constructs of pain.

Introduction

Chronic pain is a dominating symptom among patients with knee osteoarthritis (OA), and is associated with deteriorating function, e.g. walking disabilities. Pain and function are recommended as core outcome measures in clinical research on OA, and it is crucial that the assessment instruments are valid, discriminative, and sensitive to change in order to determine the effect size of a given treatment. A variety of treatment options exist, but clinical trials show minimal to modest effect sizes, and negative results for treatments previously shown efficacious. These variations may be caused by differences in study populations or methodological biases, e.g. in relation to the underlying constructs of assessment methods. Either way it has put focus on clinical research design and methods, including the need for identification of treatment response indicators.

Quantification of pain and functional disabilities implies methodological challenges, as they are subjective and variable constructs, containing several dimensions. In clinical trials, chronic musculoskeletal pain is commonly assessed retrospectively with unidimensional scales, either with 100 mm visual analogue scales (VAS) or 11, 21 or 101 point numeric rating scales (NRS), i.e. from 0–10, 0–20 or 0–100, anchored at the extremes with verbal descriptors (e.g. 0 = “no pain”, 100 = “worst pain imaginable”). VAS and NRS are found to be valid measures of pain intensity in a variety of populations, yet critics raise questions of the psychometric properties of single item measures. In knee OA, pain and function are commonly assessed with patient-reported multidimensional questionnaires, e.g. the Knee Injury and Osteoarthritis Outcome Score (KOOS). The KOOS has shown good psychometric qualities, i.e. sensitivity to change, construct validity (with SF-36), and reliability.
has showed sensitivity to change as demonstrated in a proof-of-concept walking model\(^\text{19}\) and in a recent clinical study\(^\text{20}\). The authors state that the standardisation of current pain, i.e. evoking pain with a pain provoking function, yields a more direct and unbiased response. However, walking models as assessment instruments have not been tested against a validated measurement instrument. Our objective was to explore whether pain measured with KOOS differs from pain during walking measured with a single item 0–100 mm VAS after a walking task.

**Materials and methods**

Data from knee OA patients included in the dietary intervention study, "The Influence of Weight loss or Exercise on Cartilage in Obese Knee OA Patients (The CAROT Trial)"\(^\text{22}\), were used in the current study. Eligibility criteria for the CAROT participants were: Obese, having a body–mass index (BMI) of more than 30 kg/m\(^2\); more than 50 years of age, primary knee OA diagnosed according to the American College of Rheumatology criteria\(^\text{23}\); clinical symptoms and radiographically or arthroscopically verified osteoarthritis in one or both knees; and ability to walk independently (without walking aid). The CAROT study was designed with an initial weight loss intervention period of 16 weeks, after which a maintenance period of additional 52 weeks was commenced. At inclusion into the CAROT study, the patients’ more symptomatic knee was determined and designated as the target knee throughout the study. If both knees were involved, the more symptomatic knee was chosen based on patient reports. The CAROT study was approved by the local ethical committee (1D: H-B-2007-088). The present study presents cross-sectional data recorded after the initial 16 weeks weight loss intervention. Eligibility criteria for inclusion to this particular study were complete recordings of the KOOS questionnaire and VAS scores during walking.

**Pain scoring**

As part of the CAROT study the patients underwent a biomechanical gait analysis. The full gait analysis protocol and results are given elsewhere\(^\text{24}\). In brief, the patients walked a 10 m walkway 15–20 times at a self-selected comfortable walking speed. After completion of the gait analyses, the patients were asked to rate their target knee pain during the gait analyses on a 100 mm VAS ("Please indicate your pain intensity while walking, on this line going from no pain to worst pain imaginable"), with the extremes anchored; 0 = “no pain” and 100 = “worst pain imaginable”. This score is referred to as VASwalking. The data were collected by an experienced PhD-student, who was blinded to the KOOS data.

Before the gait analysis and walking pain ratings, the KOOS\(^\text{31}\) were filled in by the patients, related to their target knee. KOOS consists of 42 items distributed on five subscales; Pain, other Symptoms, Function in daily living, Function in sport and recreation, and knee-related Quality of life. Items relate to experiences during the last week. Standardised response options are given on five-point Likert scales, and each item response equals a score from 0 to 4 (none = 0, mild = 1, moderate = 2, severe = 3, extreme = 4). A normalised score from 0 to 100 is calculated for each subscale (100 = no symptoms and 0 = extreme symptoms). In the present study, the KOOS pain subscale and item 5 in the KOOS pain subscale, KOOSp05 ("What amount of knee pain have you experienced the last week during walking on flat surface") were extracted for analyses.

**Statistical analysis**

The agreement between VASwalking and the KOOSp05 item was illustrated by a diagram showing the
distribution of VASwalking scores within the five response categories of the KOOSp05 item. To assess, any association of Spearman correlation coefficients between VASwalking and KOOSp05 was calculated. The agreement between the KOOS pain subscale and VASwalking were estimated with the Limits of Agreement method described by Bland and Altman. To ease the interpretability of the results, the normalised KOOS pain subscale score was reversed to the same polarity as the VAS score (i.e. 0 = no pain and 100 = extreme pain).

Results

Complete recordings of pain from both VASwalking and KOOS questionnaire at the 16 week follow-up session were available from 143 patients. Demographic and baseline characteristics are presented in Table 1, showing no differences between the entire CAROT cohort and the sub-cohort of participants included in this study.

The distribution of VASwalking scores within KOOSp05 response categories is seen in Figure 1, also displaying the VASwalking median within each category. On a group level, participants who rated their pain as mild (category 2) and moderate (category 3) in the questionnaire displayed uniform VAS pain ratings with medians of 10 (range 0–55) and 11 (range 0–60), respectively. It should be noted that no participants rated their pain as extreme (category 4), and only three rated their pain as severe (category 3). The VASwalking and KOOSp05 were positively associated, assessed by the Spearman correlation coefficient; \( r = 0.51; p < 0.0001 \).

The relation between the reversed KOOSpain scores and the VASwalking scores is presented in Figure 2, with line of equality. Mean VASwalking score was 11.1 mm (SD 13.3 mm), and mean KOOSpain score was 70.0 (SD 18.6), making the mean reversed KOOSpain score 30.0 (KOOSrev). Means are presented in Table 1. Mean difference between VASwalking and KOOSpain scores was 18.8 (SD 16.6, 95% confidence interval 16.1 to 21.6), and the Limits of Agreement (mean + 1.96SD) were –13.6 and 51.3, as illustrated in Figure 3.

Discussion

This large sample reveals disagreement between methods intended to measure a construct of OA-related knee pain. The poor agreement between KOOS pain subscale and VASwalking is not surprising given that the KOOS subscale contains more aspects of pain than present during walking. This may also explain why VAS scores are generally lower than KOOS scores in this population (mean difference in pain scores: 18.8), and indicates that patients are capable of separating different aspects of pain. However, assessment of the specific construct of pain – walking on a flat surface – yields divergent scores, as reflected in the distribution of VASwalking scores within the KOOSp05 item response categories, and the relatively modest correlation between the scores. Similar results are presented in the study of a walking model to assess the onset of analgesia in knee OA patients, where the time weighted average pain intensity, rated on an 11-point NRS during walking was compared to ratings on the Western Ontario Mcmaster Universities Osteoarthritis Index.

Table 1 Participants characteristics. \( p \) values are two tailed, unpaired with different variance

<table>
<thead>
<tr>
<th></th>
<th>CAROT (n 173)</th>
<th>Not eligible (n 30)</th>
<th>Study sample (n 143)</th>
<th>( p ) value of CAROT and study sample</th>
<th>( p ) value of Not eligible and study sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>Mean (SD)</td>
<td>62.99 (6.33)</td>
<td>65.6 (5.7)</td>
<td>62.45 (6.3)</td>
<td>.453</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.01</td>
</tr>
<tr>
<td>Range</td>
<td>50.1 – 76.9</td>
<td>53.3 – 73.3</td>
<td>50.1 – 76.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females n (%)</td>
<td>140 (81.03)</td>
<td>23 (76.7)</td>
<td>117 (81.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males n (%)</td>
<td>33 (18.97)</td>
<td>7 (23.3)</td>
<td>26 (18.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>Mean (SD)</td>
<td>31.95 (4.13)</td>
<td>33.3 (4.4)</td>
<td>31.67 (4.0)</td>
<td>.541</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.31</td>
</tr>
<tr>
<td>Range</td>
<td>24.8 – 47.1</td>
<td>26.6 – 47.1</td>
<td>24.8 – 45.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOOS pain subscale*</td>
<td>Mean (SD)</td>
<td>68.24 (18.93)</td>
<td>59.72 (18.36)</td>
<td>70.03 (18.62)</td>
<td>.400</td>
</tr>
<tr>
<td>VAS pain after walking**</td>
<td>Mean (SD)</td>
<td>-</td>
<td>-</td>
<td>11.13 (13.30)</td>
<td>-</td>
</tr>
</tbody>
</table>

*0–100, 100 = no pain
**0–100, 0 = no pain

Licensee OA Publishing London 2013. Creative Commons Attribution License (CC-BY)

are measured, is not clear. Obviously there are paramount differences between a questionnaire and a performance test. An assessment method is carefully developed to grasp specific constructs of interest and its psychometric properties can be estimated through testing, as has been the case with KOOS and VAS. Rating pain subsequent to a walking test as presented in this study is not a standardised assessment method, and for that reason alone, comparing it against a validated questionnaire as KOOS implies interpretational limitations. Several parameters potentially influence the process of rating subjective variables as pain and function, e.g. the setting (sitting down filling out the questionnaire versus standing/sitting in a gait laboratory shortly after a walking task), the visual presentation (rating 42 items on a 5 point Likert scale in a four-page questionnaire versus marking a point on a 100 mm straight line with the anchors “no pain” and “worst pain imaginable”), and length of required retrospectivity (up to one week versus few minutes). These parameters are likely to determine the responses to some degree, as they influence which associations the participants get when presented with the items, e.g. do the responders relate to the content, are they distracted when answering, are they influenced by presence of a researcher, do they remember their walking pain/difficulties specifically or a more overall pain/disability experience? Thus, methodological biases cannot be ruled out in this study, yet the possibility of different constructs measured as an explanation of the divergent results must also be considered.

The multidimensionality of pain, containing both the sensory intensity of pain and the emotional pain perception, is reflected in the different modalities of pain treatment, i.e. pharmacological and physical versus psychosocial and behavioural treatment modalities. Self report

(WOMAC), which is included in the KOOS questionnaire. Correlation coefficients between the two outcomes were low; $r = 0.4$ for WOMAC item, “What amount of knee pain have you experienced the last 48 hours during walking on flat surface?”, $r = 0.2$ ($P = 0.01$) for WOMAC pain subscale, and $r = 0.2$ ($P = 0.08$) for WOMAC physical function subscale.

Whether these divergences between questionnaires and walking models reflect methodological conditions or that separate constructs

Figure 1: Distribution of VAS scores (0–100) within KOOS response categories (0–4) for item 5 of the pain subscale (“What amount of knee pain have you experienced the last week during walking on flat surface?”). The Spearman correlation between VAS and KOOS item scores is $r = 0.51$ ($p < 0.0001$).

Figure 2: Pain measured with KOOS and VAS, with line of equality. KOOS is reversed to the same polarity as VAS (KOOSrev).

Licensee OA Publishing London 2013. Creative Commons Attribution License (CC-BY)

Difference against mean for Pain data by KOOS and VAS. Mean Difference = 18.8, Limits of Agreement = -13.6–51.3. KOOS is reversed to the same polarity as VAS (KOOSrev).

Figure 3: Difference against mean for Pain data by KOOS and VAS. Mean Difference = 18.8, Limits of Agreement = -13.6–51.3. KOOS is reversed to the same polarity as VAS (KOOSrev).

Research study

questionnaires assess pain experience, a construct influenced by numerous known and unknown factors, physiological as well as psychosocial and contextual, i.e. embracing both the sensory and emotional components of pain. The extent to which each of these factors contributes to the pain experience is individual and variable, thus preventing a precise description of the target construct of the assessment, which is disturbing for the sensitivity of the measurement instrument and the interpretation of the results.

Analysis of construct validity conclude that scores on unidimensional numerical pain intensity rating scales reflect the emotional qualities of pain much more than its sensory intensity in post-operative, and chronic cancer patients. In contrast, pain intensity scales (NRS and VAS) seem to reflect sensory rather than emotional pain dimensions in a study on women with chronic musculoskeletal pain. The weighting of the sensory and emotional aspects of pain in OA pain reports is unknown, though it has been shown that depression and anxiety associated with chronic pain, yield higher pain scores on unidimensional scales, indicating that the emotional proportion of the pain experience increases with affective suffering. In accordance, it has been suggested that chronic pain patients have an information processing bias towards affective pain descriptors, and that chronic pain can lead to a lack of dissociation between the affective and sensory components of pain. Thus, uncertainty remains about the sensory and emotional contributions to pain experience in knee OA and thereby which constructs are measured by existing OA pain assessment methods, compromising the interpretation of results to determine treatment effects. As different treatment modalities target different components of the pain experience, the assessment methods should reflect constructs of the corresponding pain components in order to fulfill their role as response indicators. Currently, this does not seem to be the case. The generally modest effect sizes seen in chronic pain treatment trials and OA clinical trials, where patient reported outcomes are standard, cause researchers to state that there is “a crucial need for identification of indicators of treatment response”. A possible lack of construct validity of commonly used assessment methods might be a part of the problem.

It is shown that patients incorporate various factors when rating pain, e.g. impact on activities, level of distress, and comparison with “worst” and “usual” pain, but not consistently the same way, i.e. response shift. Experiencing pain and functional limitations simultaneously with pain rating may help focus on the knee pain and to a certain extent omit other factors contributing to the pain experience, as concluded in a recent study. A possible consequence of this could be a shift in the construct measured as well, i.e. while the KOOS pain experience is a broad psychosocial and contextual construct, the walking-related pain experience might be a primarily sensory construct, dominated by the physiological pain perception. Supportive to this assumption, the walking model with NRS pain intensity scoring by Peeva et al. showed sensitivity to change with relatively large effect sizes of 0.9–1.2, thus suggesting that OA knee pain can be adequately measured during walking, possibly reflecting a construct that is closer related to the target of pharmacological and training interventions, than the constructs captured with conventional questionnaires.

Licensee OA Publishing London 2013. Creative Commons Attribution License (CC-BY)

Conclusion
This study reveals poor agreement between assessments of OA-related pain by KOOS and VAS pain score subsequent to a walking task, and moderate correlation of “pain during walking” scores between a specific item of the KOOS and VAS pain score subsequent to a walking task, in this sample of knee OA patients. This suggests that different constructs are measured with the two methods, although further validation studies are required to clarify this issue.

Acknowledgements
All authors have read and concur with the content in the final manuscript. This study was financially supported by grants from The Oak Foundation, The Velux Foundation, The Cambridge Health and Weight plan UK, The Danish Rheumatism Association, The Augustinus Foundation, The A.P. Møller Foundation for the Advancement of Medical Science, Hørslev Fonden, Bjarne Jensens Fond and Aase og Ejnar Danielsens fond. None of the sponsors had any involvement in the design of the study, in the collection, analysis, and interpretation of data, in the writing of the manuscript, or in the decision to submit the manuscript for publication.

Funding
This study was supported by grants from The Oak Foundation, The Velux Foundation, The Cambridge Health and Weight plan UK, The Danish Rheumatism Association, The Augustinus Foundation, The A.P. Møller Foundation for the Advancement of Medical Science, Hørslev Fonden, Bjarne Jensens Fond and Aase og Ejnar Danielsens fond.

Abbreviations list
BMI, body–mass index; KOOS, Knee injury and Osteoarthritis Outcome Score; NRS, Numeric Rating Scales; OA, osteoarthritis; VAS, visual analogue scale; WOMAC, Western Ontario McMaster Universities Osteoarthritis Index.

References


