Associations between single-question Visual Analogue Scale pain score and weight-bearing and non-weight-bearing domains of Western Ontario and McMaster Universities Arthritis Index pain: data from 2 phase 3 clinical trials

Jonathan J. Bjerre-Bastos\textsuperscript{a,b}, Claire P. Miller\textsuperscript{b,*}, Yanqi Li\textsuperscript{b,c}, Jeppe R. Andersen\textsuperscript{b}, Morten Karsdal\textsuperscript{d}, Asger R. Bihlet\textsuperscript{b}

Abstract

Introduction: Visual Analogue Scale (VAS) and the pain subscale of the Western Ontario and McMaster Universities Arthritis Index (WOMAC) are commonly used measuring tools of osteoarthritis (OA) pain.

Objectives: The objective of this cross-sectional study was to explore the associations between single-question VAS pain and the weight-bearing and non-weight-bearing domains of WOMAC pain.

Methods: Data from 2093 patients with OA participating in 2 phase 3 clinical trials were included for post hoc analyses. Univariate Pearson correlations and comparison of \( r \) values were made using \( z \) statistics obtained using the Fisher \( r \) to \( z \) test for all items of the VAS pain scale, the WOMAC pain subscale, the weight-bearing and non-weight-bearing constructs of WOMAC pain subscale, and by subgroups of WOMAC pain quintiles and Kellgren–Lawrence grades.

Results: The correlations between VAS pain and WOMAC pain were significant (\( r = 0.67, P < 0.001 \)) with a slope of 0.57 (95\% confidence interval [CI]: 0.54–0.61). A similar correlation was found for weight-bearing pain (\( r = 0.68, P < 0.001 \), slope: 0.62 (95\% CI: 0.59–0.65)) but significantly lower for non-weight-bearing pain (\( r = 0.55, P < 0.001 \), slope: 0.49 (95\% CI: 0.46–0.52)). The degree of disagreement between the 2 instruments seemed to be lesser in the extreme ends of the scales, and the observed association between Kellgren–Lawrence grade and disagreement between VAS and WOMAC was driven by non-weight-bearing pain.

Conclusion: In conclusion, VAS pain and WOMAC pain subscale correlation was found to be moderate and the VAS pain scale correlated more accurately with the WOMAC pain weight-bearing questions. This constitutes novel insight into patient with OA pain reporting.

Keywords: Osteoarthritis, VAS, WOMAC, Weight bearing, PROM, Pain

1. Introduction

Pain assessment in osteoarthritis (OA) clinical research commonly includes the generic unidimensional Visual Analogue Scale (VAS) as well as the multidimensional pain domain of the osteoarthritis-specific Western Ontario and McMaster Universities Arthritis Index (WOMAC).\textsuperscript{3}

Understanding the context in which a patient’s pain is reported and how various pain measurements correlate is necessary for an appropriate study design and to effectively capture drug responsiveness and perform cross-study comparisons. Some of the challenges faced by OA drug developers include lack of clarity on OA pain mechanisms\textsuperscript{10} as well as trial placebo responses, making detection of efficacy difficult.\textsuperscript{17} Despite advances in OA pain pathology and phenotyping,\textsuperscript{10} it is still largely unexplained why some patients with mild structural disease report unproportionally high levels of pain and vice versa.\textsuperscript{7} Attempts to mitigate the impact of this phenomenon in clinical research have been proposed at a protocol level and through a statistical approach, for example, by training patients in pain reporting\textsuperscript{16} and statistical modelling taking into account “regression to the mean”.\textsuperscript{6} Initiatives such as the Osteoarthritis Research Society International–Outcome Measure in Rheumatology (OARSI–OMERACT) have aimed to standardize and more effectively
capture various elements of OA pain. Nonetheless, the WOMAC pain subscale and single-question OA pain assessment (VAS, Numeric Rating Scale [NRS], and Likert) continue to be used as outcome measures in clinical OA trials.

By virtue of inherent differences, the VAS pain and WOMAC pain subscale may potentially measure different entities of pain. VAS pain, derived from the NRS and first published in 1966, is considered a measurement of global pain intensity. The WOMAC pain subscale, however, is context specific on pain when walking, stair climbing, resting, sleeping, and standing. Knowledge of the correspondence, or lack thereof, between these commonly used scales is important for the development of clinical trial protocols, where pain scores are used for patient selection as well as for assessments of clinically meaningful reduction of pain and thus efficacy of treatment. The assumption that the score of a 0 to 100 standardized WOMAC pain score is directly interchangeable with that of a 0 to 100 standardized VAS pain score may not be accurate and could potentially lead to a skewed interpretation of outcome or an unnecessarily strict or lax pain selection criterion. Ideally, identifying scale correlations could also help better understand the type of pain a given trial patient is referring to, when asked to rate their OA pain as a single score.

To the best of our knowledge, no study has investigated the VAS/WOMAC scale correspondence nor how the different categories of weight-bearing and non–weight-bearing questions of WOMAC pain subscale relate the single VAS question on general pain.

Based on the assumption that mechanisms behind weight-bearing and non–weight-bearing pain may be different and therefore not equally suitable for a general OA pain assessment, the objective of this large post hoc cross-sectional study was to investigate how 2 commonly used OA pain-reporting tools are correlated by exploring the correspondence and degree of agreement between the single-question VAS pain score and individual dimensions of the WOMAC pain subscale.

2. Methods

2.1. Study population

Baseline data from a large clinical data set of patients with OA based on 2 phase 3 clinical trials investigating oral salmon calcitinin, CSMC021C2301 (NCT00486434) and CSMC021C2302 (NCT00704847), were analyzed post hoc. Subjects with available VAS and WOMAC pain subscale scores of the nontarget knee at baseline were included in the analysis (n = 2,093).

2.2. Single-question Visual Analogue Scale for pain

Visual Analogue Scale has been accepted as a global pain measure for decades and widely used in clinical settings to assess daily pain levels. In these 2 trials, VAS was designed as a visual analogue scale ranging from 0 to 100 mm anchored with “No pain” (0) and “Worst pain imaginable” (100). Study participants were encouraged to abstain for analgesic medication 3 days before study enrollment and hereby the collection of pain outcome scores. Prompted by the question “How was your pain level the past 24-hours?” participants were asked to mark their pain level on the line between the 2 end points. The intention being that the open-ended question did not relate to a specific activity (eg, rest or walking) but reflecting the overall average pain levels the last day as perceived by the participant.

2.3. Western Ontario and McMaster Universities Arthritis Index

The WOMAC scale is a self-administered questionnaire consisting of 24 items divided into 3 subscales: pain (5 items), stiffness (2 items), and physical function (17 items). The analysis did not include the WOMAC function or WOMAC stiffness subscales. The pain subscale consists of 5 individual questions and is considered a reference standard assessing both weight-bearing (questions 1, 2, and 5) and non–weight-bearing (questions 3 and 4) pain separately. In these 2 trials, the WOMAC scale was assessed at the same time as VAS after the analgesic wash-out period. The participants were asked to read each question in detail, assess pain during the past 24 hours, and mark an X on a 100-mm line, on which 0 mm equaled “No Pain” and 100 equaled “Extreme Pain” with the full WOMAC pain subscale score ranging from a total of 0 to 500, and in the context of the current analysis, standardized to 0 to 100.

2.4. Statistical analysis

The analyses were focused on the nontarget knee to avoid selection bias of the target knee because inclusion into the studies required a target knee score >150 of 500 on the WOMAC pain subscale (corresponding to 30 of 100). Before analyses, normality and homoscedasticity were assessed by visual inspection of normal and residuals plots. Linear regression analysis was used to calculate the relationship between the VAS pain score and the full WOMAC pain subscale score standardized to a 0 to 100 scale. In addition, correlation analyses were performed using linear regression between VAS and the respective scores of weight-bearing (questions 1, 2, and 5) and non–weight-bearing pain (questions 3 and 4), both standardized to a 0 to 100 scale. P values were calculated using Spearman correlation in case of non-normality. We also investigated correlations between the first and fifth individual quintile of VAS and the 5 individual WOMAC pain questions. To differentiate between pain constructs of WOMAC pain subscale (weight-bearing vs non–weight-bearing), the correlation coefficients of the VAS or weight-bearing pain score correlation and the VAS or non–weight-bearing pain score correlation were, respectively, compared with VAS or full pain score correlation using z statistics obtained using Fisher dependent samples r to z. Finally, we examined the possible influence of standard baseline parameters including age, sex, body mass index, and Kellgren–Lawrence (KL) grade as independent variables, on divergence of scales using multiple regression analysis. Associations between KL grade, an ordinal variable, and absolute differences between respective scores of full WOMAC pain, weight-bearing pain, and non–weight-bearing pain and VAS score were further explored using the Kruskal–Wallis test.

P values <0.05 was considered significant, although all analyses were considered exploratory and no adjustment for multiplicity was made. Statistical analyses were performed in MedCalc version 20.010 and GraphPad Prism version 9.1.0.

3. Results

3.1. Study sample

The database included 2206 study participants. Data were missing at random for 113 participants, of which 84 had incomplete or missing WOMAC pain subscale and 29 were missing VAS score for nontarget knee, resulting in 2,083 participants included for analyses. The mean age was 64.4 years...
3.2. Agreement between scales

The correlation between single-question VAS pain and full WOMAC pain score was significant ($r = 0.67$, $P < 0.001$) with a slope of 0.57 (95% confidence interval: 0.54–0.61) (Fig. 1). When dividing into 2 constructs representing weight-bearing (Q1, Q2, and Q5) and non–weight-bearing pain (Q3 and Q4), the correlations with VAS were also significant ($r = 0.68$ and 0.55, respectively, both $P < 0.001$) (Fig. 2). When compared with the VAS or full WOMAC pain score correlation, the correlation coefficient was not different for weight-bearing pain ($Fisher z = -0.73$, $P = 0.5$) but significantly lower ($Fisher z = 10.5$, $P < 0.001$) for non–weight-bearing pain. The correlations between VAS and individual WOMAC pain questions within quintiles of VAS scores were generally weak ($r < 0.35$). As summarized in Table 2, agreement between individual WOMAC pain questions was higher in the extreme ranges of VAS measured pain, ie, the first and fifth VAS quintiles. The strongest correlations were found between low VAS pain and weight-bearing WOMAC pain questions. Generally, the slopes and correlation strengths were low for WOMAC non–weight-bearing pain dimensions and the middle, second, third, and fourth VAS quintiles.

3.3. Factors associated with disagreement

There was a significant association between KL grade and VAS or WOMAC scale discordance ($P = 0.002$) (VAS score subtracted by WOMAC score, both standardized to a 0–100 value) and when subanalyzing for non–weight-bearing pain ($P < 0.001$) (Figs. 3A–C). No other baseline parameters correlated to scale discordance between VAS and WOMAC weight-bearing pain. A significant trend was found for larger mean divergence in VAS and WOMAC pain with increasing KL grade ($P = 0.72$), but a highly significant trend was found for larger mean divergence in VAS and WOMAC non–weight-bearing pain with increasing KL grade ($P < 0.0001$) (Figs. 3A–C).

4. Discussion

4.1. Summary

In this cross-sectional study of baseline data from 2,093 patients with knee OA, we explored the associations between the individual dimensions of pain-reporting instruments, single-question global VAS pain, and WOMAC pain subscale, to determine correspondence and degree of agreement between scales.

Our findings indicate a moderate correlation between VAS pain and WOMAC pain. The 95% limits of agreement between scales, measuring the presumed same outcome, were wide. On average, an increase of 1 unit in VAS pain score corresponded 0.57 unit of change (95% confidence interval: 0.54–0.61) in WOMAC pain score (standardized 0–100). However, the agreement between scales was notably increased for the weight-bearing construct of WOMAC pain, particularly at the most extreme ranges of VAS measured pain ie, the first and fifth quintiles of VAS. The strongest correlation overall was found between low VAS pain score (first quintile) and pain during stair climbing (WOMAC pain Q2, slope: 1.36, $r = 0.34$ ($P < 0.001$), consistent with “stair climbing” being the first reported symptom for patients with knee OA.15 The extent of scale discordance in absolute numbers (VAS pain minus WOMAC pain scale (standardized to 0–100) was found to increase with progression in the degree of structural severity as assessed by KL grade ($P = 0.0002$). This association seemed to be driven by non–weight-bearing pain ($P < 0.0001$) while there was no significant trend when analyzing for weight-bearing pain (Fig. 3).

4.2. Interpretation

This study indicates that when patients with OA are asked a single, less specific question, the answer might intuitively pertain to pain in weight-bearing activities such as standing, walking, and climbing stairs. This could support the notion that patients early in

![Figure 1. Correlation between VAS pain score and total WOMAC pain score. Spearman $r$ and $P$ values are reported. VAS, Visual Analogue Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index.](image-url)
OA disease progression experience nociceptive pain on weight-bearing activities. This is supported by the data indicating that the agreement between VAS and WOMAC pain weakens with increasing structural severity, primarily because of an associated increase in non-weight-bearing pain. Weight-bearing pain has been described as more likely to reflect biochemical mechanisms triggered by tissue damage leading to nociceptive pain, whereas the pathology of non-weight-bearing pain continues to be poorly understood. It has been hypothesized to arise as a neuroplastic response to the chronicity of nociceptive input from the subchondral bone marrow lesions albeit this hypothesis has been challenged in a recent experimental study suggesting that the complicated cognitive abstraction required in scoring pain in the intermediate range contributed to this finding, indicating that the complicated cognitive abstraction required in scoring pain in the intermediate range contributed to this finding, suggesting that the complicated cognitive abstraction required in scoring pain in the intermediate range contributed to this finding, although the potential explanations for this observation were not possible to study in the current data set.

### 4.3. Clinical implications

Our findings demonstrate that a standardized outcome on WOMAC pain does not directly translate 1:1 to the 24-hour VAS pain score. Although the current analysis was made in the context of VAS, the finding can be presumed to apply to other global single-question pain assessments, ie, NRS and average daily pain as well. The VAS pain score and WOMAC pain subscale operate with “worst pain imaginable” or “extreme pain” anchored at 100; however, as shown in Figure 1, worst pain imaginable on VAS pain (VAS 100) only translates to an average WOMAC pain subscale (0–100) score of approximately 72. These pain measures can thus not be used interchangeably, and in doing so, the value of pain being measured could be misinterpreted, directly or indirectly affect the outcome of a clinical trial, or lead to underestimation of a patient’s pain and insufficient pain management in a clinical setting.

### 4.4. Limitations

A potential limitation is that our findings are predominantly based on patients with OA with mild-to-moderate disease severity as 52 participants (2.5%) had a nontarget knee KL grade of 4. Yet we consider it a substantial data set for deducting correlations of severe KL grades to scale agreement because of the exploratory nature of the analyses. Finally, a potential limitation could be the inclusion and analyses of nontarget knee pain data. It cannot be ruled out that the participants are unable to neglect the symptoms from the target knee when asked to access their pain levels. However, study participants were eligible with bilateral knee OA, and selection of target knee in the case of both knees reaching the

### Table 2

| WOMAC pain Q1 (walking) | WOMAC pain Q2 (stairs) | WOMAC pain Q3 (nocturnal) | WOMAC pain Q4 (rest) | WOMAC pain Q5 (standing) |
|-------------------------|-------------------------|----------------------------|----------------------|-------------------------|
| VAS 1st quintile (n = 584) | Slope: 0.81, r = 0.31 (P < 0.001) | Slope: 1.36, r = 0.34 (P < 0.001) | Slope: 0.54, r = 0.18 (P < 0.001) | Slope: 0.74, r = 0.28 (P < 0.001) | Slope: 0.95, r = 0.33 (P < 0.001) |
| VAS 2nd quintile (n = 524) | Slope: 0.24, r = 0.07 (P = 0.09) | Slope: 0.39, r = 0.10 (P = 0.02) | Slope: 0.20, r = 0.05 (P = 0.23) | Slope: 0.25, r = 0.07 (P < 0.001) | Slope: 0.20, r = 0.06 (P < 0.001) |
| VAS 3rd quintile (n = 506) | Slope: 0.41, r = 0.11 (P = 0.01) | Slope: 0.48, r = 0.13 (P = 0.005) | Slope: 0.09, r = 0.02 (P = 0.64) | Slope: 0.34, r = 0.09 (P = 0.43) | Slope: 0.28, r = 0.07 (P = 0.10) |
| VAS 4th quintile (n = 305) | Slope: 0.44, r = 0.11 (P = 0.05) | Slope: 0.47, r = 0.14 (P = 0.01) | Slope: 0.10, r = 0.02 (P = 0.72) | Slope: 0.13, r = 0.03 (P = 0.59) | Slope: 0.21, r = 0.05 (P = 0.39) |
| VAS 5th quintile (n = 174) | Slope: 0.78, r = 0.17 (P = 0.02) | Slope: 0.76, r = 0.24 (P < 0.001) | Slope: 1.53, r = 0.28 (P < 0.001) | Slope: 1.1, r = 0.22 (P < 0.003) | Slope: 0.62, r = 0.14 (P = 0.07) |

n, number; Q, question; VAS, Visual Analogue Scale; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index. Bold entries indicate p < 0.05.
study pain criteria would be determined by KL grade, and thus, the nontarget knee included for analyses in this study is not necessarily the knee with the least pain. It would also not be expected to influence scale comparison because of the likelihood of any erroneous reporting to occur similarly on both VAS pain and WOMAC pain.

### 4.5. Conclusion

The 24-hour single-question VAS pain scale and the pain subscale of WOMAC were moderately correlated. Our results suggest that VAS pain scale correlated more accurately with the weight-bearing construct of WOMAC pain subscale as compared with the WOMAC pain non–weight-bearing construct. Agreement between VAS pain and WOMAC pain was best in the extreme ranges of the scale. Scale discordance was found to be positively correlated to the radiographic severity of OA and the observed divergence between VAS pain and WOMAC pain to be driven by the non–weight-bearing questions of WOMAC pain subscale.

### Disclosures

J.J. Bjerre-Bastos, C.P. Miller, Y. Li, J.R. Andersen, and A.R. Bihlet are full-time employees of NBCD A/S. Y. Li, J.R. Andersen, and A.R. Bihlet are shareholders of NBCD A/S. M. Karsdal is a shareholder of Nordic Bioscience A/S.

### Acknowledgements

The CSMC212301/02 trials were cosponsored by Novartis and Nordic Bioscience; however, no specific grant from funding agencies in the public, commercial, or not-for-profit sectors was received for the current report.

### Article history:

Received 1 April 2022
Received in revised form 5 May 2022
Accepted 6 May 2022

### References

[1] Aoyagi K, Liew J, Carkesso L, Famer J, Wang N, Frey Law L, Lewis C, Nevitt M, Neogi T. Weight-bearing and non-weight-bearing pain may reflect different pain mechanisms in knee osteoarthritis. Osteoarthr Cartil 2022;30:545–50.

[2] Aso K, Shahtaheri SM, McWilliams DF, Walsh DA. Association of subchondral bone marrow lesion localization with weight-bearing pain in people with knee osteoarthritis: data from the Osteoarthritis Initiative. Arthritis Res Ther 2021;23:35.

[3] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1835–40.

[4] Bihlet AR, Bjerre-Bastos JJ, Andersen JR, Byrjalsen I, Karsdal MA, Bay-Jensen AC. Clinical and biochemical factors associated with risk of total joint replacement and radiographic progression in osteoarthritis: data from two phase III clinical trials. Semin Arthritis Rheum 2020;50:1374–81.
Bihlet AR, Byrjalsen I, Bay-Jensen AC, Andersen JR, Christiansen C, Riis BJ, Karsdal MA. Associations between biomarkers of bone and cartilage turnover, gender, pain categories and radiographic severity in knee osteoarthritis. Arthritis Res Ther 2019;21:203.

Bihlet AR, Byrjalsen I, Bay-Jensen AC, Andersen JR, Christiansen C, Riis BJ, Valter I, Karsdal MA, Hochberg MC. Identification of pain categories associated with change in pain in patients receiving placebo: data from two phase 3 randomized clinical trials in symptomatic knee osteoarthritis. BMC Musculoskelet Disord 2018;19:17.

Bjerre-Bastos JJ, Byrjalsen I, Karsdal MA, Andersen JR, Bihlet A. SAT0634 co-morbidities associated with discordance between structural severity and pain in osteoarthritis: implications for clinical trial design in OA—a post-hoc analysis of data from two randomized controlled trials. Ann Rheum Dis 2019;78:1414–1415.

Bond MR, Pilowsky I. Subjective assessment of pain and its relationship to the administration of analgesics in patients with advanced cancer. J Psychosom Res 1966;10:203–8.

Conaghan PG, Hunter DJ, Cohen SB, Kraus VB, Berenbaum F, Lieberman JR, Jones DG, Spitzer AI, Jevsevar DS, Katz NP, Burgess DJ, Bodick N. Effects of a single intra-articular injection of a microsphere formulation of triamcinolone acetonide on knee osteoarthritis pain: a double-blinded, randomized, placebo-controlled, multinational study. J Bone Jt Surg Am Vol 2018;100:666–77.

Deveza LA, Nelson AE, Looser RF. Phenotypes of osteoarthritis: current state and future implications. Clin Exp Rheumatol 2019;37(suppl 120):64–72.

Fu K, Robbins SR, McDougall JJ. Osteoarthritis: the genesis of pain. Rheumatology 2018;57:iv43–iv50.

Hawker GA, Davis AM, French MR, Cibere J, Jordan JM, March L, Suarez-Almazor M, Katz JN, Dieppe P. Development and preliminary psychometric testing of a new OA pain measure—an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:409–14.

Hensor EMA, Dube B, Kingsbury SR, Tennant A, Conaghan PG. Toward a clinical definition of early osteoarthritis: onset of patient-reported knee pain begins on stairs. Data from the osteoarthritis initiative. Arthritis Care Res 2015;67:40–7.

Hjemstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, Fainsinger R, Aass N, Kaasa S. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. J Pain Symptom Manage 2011;41:1073–93.

McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, Wang S, Schnitzer T. OARSI Clinical Trials Recommendations: design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage 2015;23:747–60.

Treister R, Lawal OD, Shecter JD, Khurana N, Bothmer J, Field M, Harte SE, Kruger GH, Katz NP. Accurate pain reporting training diminishes the placebo response: results from a randomised, double-blind, crossover trial. PLoS One 2018;13:e0197844.

Zhang W. The powerful placebo effect in osteoarthritis. Clin Exp Rheumatol 2019;37(suppl 120):118–23.