Journal of Pain Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Test–retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain

This article was published in the following Dove Press journal:
Journal of Pain Research

Ahmad H Alghadir
Shahnawaz Anwer
Amir Iqbal
Zaheen Ahmed Iqbal

Rehabilitation Research Chair, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

Objective: Several scales are commonly used for assessing pain intensity. Among them, the numerical rating scale (NRS), visual analog scale (VAS), and verbal rating scale (VRS) are often used in clinical practice. However, no study has performed psychometric analyses of their reliability and validity in the measurement of osteoarthritic (OA) pain. Therefore, the present study examined the test–retest reliability, validity, and minimum detectable change (MDC) of the VAS, NRS, and VRS for the measurement of OA knee pain. In addition, the correlations of VAS, NRS, and VRS with demographic variables were evaluated.

Methods: The study included 121 subjects (65 women, 56 men; aged 40–80 years) with OA of the knee. Test–retest reliability of the VAS, NRS, and VRS was assessed during two consecutive visits in a 24 h interval. The validity was tested using Pearson's correlation coefficients between the baseline scores of VAS, NRS, and VRS and the demographic variables (age, body mass index [BMI], sex, and OA grade). The standard error of measurement (SEM) and the MDC were calculated to assess statistically meaningful changes.

Results: The intraclass correlation coefficients of the VAS, NRS, and VRS were 0.97, 0.95, and 0.93, respectively. VAS, NRS, and VRS were significantly related to demographic variables (age, BMI, sex, and OA grade). The SEM of VAS, NRS, and VRS was 0.03, 0.48, and 0.21, respectively. The MDC of VAS, NRS, and VRS was 0.08, 1.33, and 0.58, respectively.

Conclusion: All the three scales had excellent test–retest reliability. However, the VAS was the most reliable, with the smallest errors in the measurement of OA knee pain.

Keywords: visual analog scale, numerical rating scale, verbal rating scale, pain, osteoarthritis, knee

Introduction

Knee pain is the major symptom of OA among the elderly, affecting both men and women.[1] Assessment of pain intensity is one of the primary outcomes used to determine the progression of OA.[2–6] Several scales are commonly used for the assessment of pain intensity.[7,8] Among them, the NRS, VAS, and VRS are often used in clinical practice.[7] These pain-rating scales have shown good validity and reliability for assessing pain intensity; however, none has shown superiority over the others,[7,8] as various aspects, such as the response categories, patient preference, application methods, and correction for missing information, make each of them unique.[9]

Because it is easy to understand and administer, the NRS is preferred over the VAS by the elderly population.[8,10] For instance, the administration of VAS requires a patient

Correspondence: Shahnawaz Anwer
Rehabilitation Research Chair, College of Applied Medical Sciences, King Saud University, King Abdullah Road, Riyadh, Saudi Arabia
Tel +966 5 9566 8288
Email anwer_shahnawazphysio@rediffmail.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/JPR.S158847

to perceive his or her pain level mathematically, which could be particularly difficult[11,12] for the elderly. In addition, the VAS has a higher reported failure rate of 7–16% when compared to the NRS and VRS.[13–15] Although the NRS and VRS are well correlated with and sensitive for pain assessment, the NRS shows a higher reliability, specifically in elderly and less educated patients, and is useful for the assessment of chronic pain.[16–18] Furthermore, a previous study showed higher responsiveness of the NRS compared to the VRS for the assessment of chronic pain.[19]

A previous study indicated low intra-scale agreement between the VAS and VRS, and recommended against their interchangeable use for musculoskeletal pain.[20] Another study reported that the three scales are sensitive for the assessment of chronic OA pain, with no differences among them.[21] Similarly, Bellamy et al[22] reported that all scales for pain measurement are able to detect statistically significant and clinically important improvements in OA pain following a pharmacological intervention. However, these studies did not perform psychometric analyses of their reliability and validity for the assessment of OA pain. Therefore, the present study examined the test–retest reliability and MDC of the VAS, NRS, and VRS for the measurement of OA knee pain.

Methods

Ethics approval

Ethical approval was obtained from the IRB, Rehabilitation Research Chair, King Saud University. In accordance with the IRB guidelines of the institution, written informed consent was obtained from each individual who agreed to participate.

Participants

The participants (n = 121) included 65 women and 56 men aged 40–80 years with OA of the knee. The severity of OA of the knee was assessed using the K–L scale.[23] Patients with any history of neurological disease, neuropathic pain, inflammatory joint disease, diabetes, or inability to cooperate were excluded.

Pain measures

The VAS is a reliable, valid, responsive, and frequently used pain outcome measure.[7] It consists of a bidirectional 10 cm straight line with two labels, that is, "no pain" and "worst possible pain", located at either end of the line. Patients are instructed to draw a vertical mark on the line indicating their pain level.[7] The NRS is an 11-point scale comprising a number from 0 through 10; 0 indicates "no pain", and 10 indicates the "worst imaginable pain". Patients are instructed to choose a single number from the scale that best indicates their level of pain.[7] The VRS is a valid scale consisting of a list of descriptors used to represent various levels of pain, including none, mild, moderate, and severe.[8,24] Patients are instructed to select the one descriptor that best indicates their pain level.

Procedure

Brief instructions about the procedure were given to the patients before they completed the scales. Demographic information such as sex, age, weight, height, BMI, and the severity of knee OA was recorded. Patients were asked to rate their baseline level of pain on the VAS,[25,26] NRS,[25,26] and VRS[25] in a random order. Patients were then requested to complete a second round of pain assessment using the VAS, NRS, and VRS after 24 h to determine test–retest reliability. Two independent examiners were involved in both sessions.

Statistical analysis

SPSS (Statistical Package for the Social Sciences) for Windows (version 22; IBM Inc., Armonk, NY, USA) was used for data analysis. Using the Shapiro–Wilk test, it was determined that the data were not normally distributed ($P < 0.05$). The test–retest reliability of VAS, NRS, and VRS was assessed using $ICC_{2,1}$. The Bland–Altman plot method was used to assess the agreement between two readings (Figures 1 and 2). The validity was tested using Pearson's correlation coefficients between the baseline scores of VAS, NRS, and VRS and the demographic variables (age, BMI,

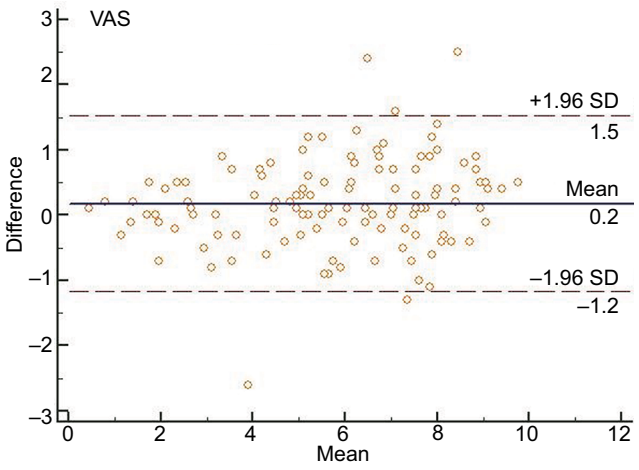

Figure 1 Bland–Altman plot: intraindividual differences (n = 121) between the visual analog scale (VAS) on test and retest, plotted against the average of the two scores.
Note: The central line represents the mean difference, and the dashed lines display the 95% limit of agreement.

sex, and OA grade). The SEM and MDC were calculated based on the results of reliability analyses.[27–29] For every test, the level of significance was set at $P < 0.05$ with a 95% confidence interval.

Results

Table 1 presents the characteristics of the participants. The average age was 52.9 years. There were 65 women and 56 men. Most of the participants had a K–L grade of 2–3 (75.3%). Baseline VAS, NRS, and VRS scores are presented in Table 2.

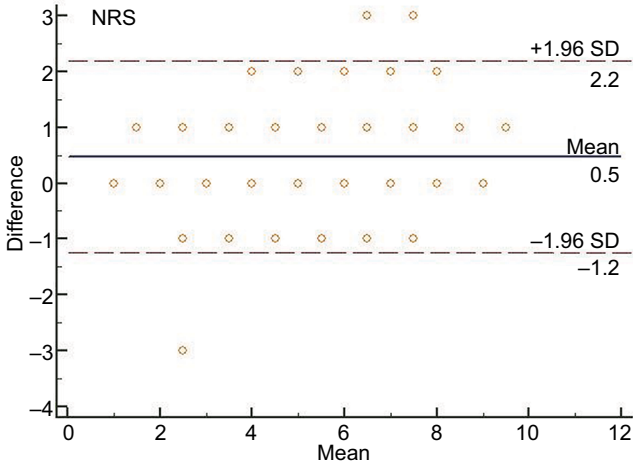

Figure 2 Bland–Altman plot: intraindividual differences (n = 121) between the numerical rating scale (NRS) on test and retest, plotted against the average of the two scores.
Note: The central line represents the mean difference, and the dashed lines display the 95% limit of agreement.

Table 1 Participants' characteristics

Characteristics	All participants
Gender, n (%)	
Male	56 (46.3)
Female	65 (53.7)
Age (years)	
Mean (SD)	52.9 (12.5)
Range	40–80
Height (m)	
Mean (SD)	1.7 (0.06)
Range	1.5–1.9
Weight (kg)	
Mean (SD)	86.4 (13.9)
Range	56–123
BMI (kg/m^2)	
Mean (SD)	30.01 (4.2)
Range	21.4–41.4
K–L scale score, n (%)	
Grade 1	19 (15.7)
Grade 2	55 (45.5)
Grade 3	36 (29.8)
Grade 4	11 (9.1)

Abbreviations: BMI, body mass index; K–L, Kellgren and Lawrence.

Test–retest reliability, SEM, and MDC

The reliability data are presented in Table 3. The ICC of the VAS, NRS, and VRS was 0.97, 0.95, and 0.93, respectively. The Bland–Altman plot showed reasonable agreement between the test–retest scores of the VAS and NRS (Figures 1 and 2, respectively). The SEM of VAS, NRS, and VRS was 0.03, 0.48, and 0.21, respectively. The MDC of VAS, NRS, and VRS was 0.08, 1.33, and 0.58, respectively (Table 3).

Validity

A good-to-excellent correlation was found between the VAS and NRS ($r = 0.941$), VAS and VRS ($r = 0.878$), and NRS and VRS scores ($r = 0.925$). Age, BMI, and OA grades were significantly correlated with all the three scales (Table 4). The sex of the participants did not correlate with any scale (Table 4).

Table 2 Descriptive statistics of baseline scores

Variables	All participants
VAS (0–10 cm)	
Mean (SD)	5.8 (2.2)
Range	0.5–10
NRS (0–10)	
Mean (SD)	5.9 (2.2)
Range	1–10
VRS n [%]	
Mild	30 (24.8)
Moderate	41 (33.9)
Severe	50 (41.3)

Abbreviations: VAS, visual analog scale; NRS, numerical rating scale; VRS, verbal rating scale.

Table 3 Reliability data of VAS, NRS, and VRS

	ICC (95% CI)	SEM	MDC
VAS	0.97 (0.96–0.98)	0.03	0.08
NRS	0.95 (0.93–0.96)	0.48	1.33
VRS	0.93 (0.90–0.95)	0.21	0.58

Abbreviations: VAS, visual analog scale; NRS, numerical rating scale; VRS, verbal rating scale; ICC, intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement; MDC, minimal detectable change.

Table 4 Correlation of VAS, NRS, and VRS with demographic variables

	VAS	NRS	VRS
Age	0.262*	0.224*	0.261*
BMI	0.379**	0.359**	0.399**
Gender	0.071	0.056	0.048
K–L grade	0.844**	0.817**	0.754**
NRS	0.941**	–	
VRS	0.878**	0.925**	–

Notes: *Correlation, $P < 0.05$. **Correlation, $P < 0.001$.
Abbreviations: VAS, visual analog scale; NRS, numerical rating scale; VRS, verbal rating scale; BMI, body mass index; K–L, Kellgren and Lawrence.

Discussion

The present study assessed the relative and absolute reliability, validity, and MDC of the three pain scales (VAS, NRS, and VRS) for the measurement of OA knee pain. The relative reliability, measured by the ICC, was excellent for all the three scales. The VAS was the most reliable (ICC = 0.97), and the VRS was the least reliable (ICC = 0.93). To the best of our knowledge, no prior study has assessed the reliability of VAS, NRS, and VRS for the measurement of OA knee pain; thus, a direct comparison of the present findings with those from other studies could not be made. One study reported excellent reliability for VAS and NRS (0.97 and 0.99, respectively) for the measurement of musculoskeletal pain.[16] Ferraz et al[30] reported that the three scales had excellent reliability when used in literate and illiterate groups of patients with rheumatoid arthritis. However, in the former study, the sample comprised a group of less educated orthopedic patients. In contrast, the present study did not consider the level of education of the patients with OA of the knee.

A few previous studies compared different pain-rating scales in patients with OA of the knee.[21,22,31] Bashir et al[21] reported good sensitivity of VAS, NRS, and VRS for assessment of pain in chronic OA; however, there were no significant differences. In addition, Bellamy et al[22] reported that all measures of pain were able to detect statistically significant and clinically important improvements in OA pain following a pharmacological intervention. However, these studies did not report the validity and reliability of these three scales in patients with OA of the knee.

Although ICC provides a good measure of reproducibility, it was found to be sensitive to heterogeneity in the sample studied.[32] On the contrary, SEM provides a more reliable and practical indicator, and helps calculate the MDC.[32] In the current study, among the three scales, the VAS had minimal measurement error (SEM = 0.03). The MDC of the NRS (MDC = 1.33 points) was higher when compared to that of the VAS and VRS. In a previous study, Childs et al[33] reported a slightly higher MDC value (MDC = 2 points) for NRS in patients with low-back pain. However, to our knowledge, no study has investigated the SEM and MDC for VAS, NRS, and VRS in patients with OA of the knee.

A good-to-excellent correlation was observed for VAS, NRS, and VRS, supporting the consistency of pain measurements among the three scales in patients with OA of the knee. These findings indicate that all the three scales are valid in assessing pain levels in patients with OA of the knee. Previous studies indicated a high correlation between VAS and NRS in patients with rheumatoid arthritis and chronic pain.[29,34] In addition, the current study indicated a good correlation between VAS and VRS ($r = 0.878$). Similarly, Bolognese et al[35] and Averbuch and Katzper[36] reported a good correlation between the VAS and categorical pain scales in OA of the knee. Moreover, in the current study, VAS, NRS, and VRS correlated well with the radiographic severity of OA of the knee as measured using the K–L grading scale. This finding supports the predictive ability of these three scales to measure the radiographic severity of OA of the knee, but this needs to be confirmed by a future study.

The current study has some potential limitations that need to be considered while interpreting these findings. The study results were limited only to patients with pain due to OA of the knee; therefore, the application of these scales to other causes of musculoskeletal pain requires some caution. In the current study, the participants had a history of pain for more than 6 months, indicating chronic pain. Thus, additional research is required to assess the use of these three scales in populations with different stages of pain. Furthermore, in the current study, patients' education level was not assessed. It is assumed that the educational status of the patients might affect pain perception; therefore, the educational background should be assessed for the effect on these scales.

Conclusion

All the three scales showed excellent test–retest reliability. A good-to-excellent correlation was found among the three scales. The VAS, NRS, and VRS were significantly correlated with the demographic variables. However, the VAS was considered the most stable, with the smallest error in measurement and MDC values for OA knee pain. Long-term follow-up studies are required to determine the responsiveness of these three scales in patients with OA of the knee.

Abbreviations

BMI, body mass index; ICC, intraclass correlation coefficient; IRB, institutional review board; K–L, Kellgren and Lawrence; MDC, minimum detectable change; NRS, numerical rating scale; OA, osteoarthritis; SEM, standard error of measurement; VAS, visual analog scale; VRS, verbal rating scale.

Acknowledgments

The authors are grateful to the Deanship of Scientific Research, King Saud University, for funding through Vice Deanship of Scientific Research Chairs. They thank the

Deanship of Scientific Research and RSSU at King Saud University for their technical support.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Jan MH, Lin CH, Lin YF, Lin JJ, Lin DH. Effects of weight-bearing versus nonweight-bearing exercise on function, walking speed, and position sense in participants with knee osteoarthritis: a randomized controlled trial. *Arch Phys Med Rehabil*. 2009;90(6):897–904.
2. Thomas KS, Muir KR, Doherty M, Jones AC, O'Reilly SC, Bassey EJ. Home based exercise program for knee pain and knee osteoarthritis: randomized controlled trial. *Br Med J*. 2002;325(7367):752–757.
3. Bruce-Brand RA, Walls RJ, Ong JC, Emerson BS, O'Byrne JM, Moyna NM. Effects of home-based resistance training and neuromuscular electrical stimulation in knee osteoarthritis: a randomized controlled trial. *BMC Musculoskelet Disord*. 2012;13:118.
4. Sled EA, Khoja L, Deluzio KJ, Olney SJ, Culham EG. Effect of a home program of hip abductor exercises on knee joint loading, strength, function, and pain in people with knee osteoarthritis: a clinical trial. *Phys Ther*. 2010;90(6):895–904.
5. Chaipinyo K, Karoonsupcharoen O. No difference between home-based strength training and home-based balance training on pain in patients with knee osteoarthritis: a randomized trial. *Aust J Physiother*. 2009;55(1):25–30.
6. Bezalel T, Carmeli E, Katz-Leurer M. The effect of a group education programme on pain and function through knowledge acquisition and home-based exercise among patients with knee osteoarthritis: a parallel randomized single-blind clinical trial. *Physiotherapy*. 2010;96(2):137–143.
7. Hjermstad MJ, Fayers PM, Haugen DF, et al. 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(6):1073–1093.
8. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. New York: Guilford Press; 2011:19–44.
9. Dworkin RH, Turk DC, Farrar JT, et al; IMMPACT. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1–2):9–19.
10. Gagliese L. Assessment of pain in elderly people. In: Turk DC, Melzack R, editors. *Handbook of Pain Assessment*. New York: Guilford Press; 2002:119–133.
11. Ohnhaus EE, Adler R. Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. *Pain*. 1975;1(4):379–384.
12. Woodforde JM, Merskey H. Some relationships between subjective measures of pain. *J Psychosom Res*. 1972;16(3):173–178.
13. Chanques G, Viel E, Constantin JM, et al. The measurement of pain in intensive care unit: comparison of 5 self-report intensity scales. *Pain*. 2010;151(3):711–721.
14. Dijkers M. Comparing quantification of pain severity by verbal rating and numeric rating scales. *J Spinal Cord Med*. 2010;33(3):232–242.
15. Hartrick CT, Kovan JP, Shapiro S. The numeric rating scale for clinical pain measurement: a ratio measure? *Pain Pract*. 2003;3(4):310–316.
16. Gallasch CH, Alexandre NM. The measurement of musculoskeletal pain intensity: a comparison of four methods. *Rev Gaucha Enferm*. 2007;28(2):260–265.
17. Taylor LJ, Harris J, Epps CD, Herr K. Psychometric evaluation of selected pain intensity scales for use with cognitively impaired and cognitively intact older adults. *Rehabil Nurs*. 2005;30(2):55–61.
18. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94(2):149–158.
19. Chien CW, Bagraith KS, Khan A, Deen M, Strong J. Comparative responsiveness of verbal and numerical rating scales to measure pain intensity in patients with chronic pain. *J Pain*. 2013;14(12):1653–1662.
20. Lund I, Lundeberg T, Sandberg L, Budh CN, Kowalski J, Svensson E. Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. *BMC Med Res Methodol*. 2005;5:31.
21. Bashir MS, Khade A, Borkar P, Saleem M, Lingaswamy V, Reddy D. A comparative study between different pain rating scales in patients of osteoarthritis. *Indian J Physiol Pharmacol*. 2013;57(2):205–208.
22. Bellamy N, Campbell J, Syrotuik J. Comparative study of self-rating pain scales in osteoarthritis patients. *Curr Med Res Opin*. 1999;15(2):113–119.
23. Kellgren JH, Lawrence JS. Radiologic assessment of osteoarthritis. *Ann Rheum Dis*. 1957;16(4):494–502.
24. Seymour RA. The use of pain scales in assessing the efficacy of analgesics on postoperative dental pain. *Eur J Clin Pharmacol*. 1982;23(5):441–444.
25. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs*. 2005;14(7):798–804.
26. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 11:S240–S252.
27. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res*. 1993;2(3):221–226.
28. Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. *J Eval Clin Pract*. 2000;6(1):39–49.
29. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 2nd ed. Norwalk: Appleton & Lange; 2000.
30. Ferraz MB, Quaresma MR, Aquino LRL, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of illiterate patients with rheumatoid arthritis. *J Rheumatol*. 1990;17(8):1022–1024.
31. Allen KD, Coffman CJ, Golightly YM, Stechuchak KM, Voils CI, Keefe FJ. Comparison of pain measures among patients with osteoarthritis. *J Pain*. 2010;11(6):522–527.
32. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30(1):1–15.
33. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine (Phila Pa 1976)*. 2005;30(11):1331–1334.
34. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis*. 1978;37(4):378–381.
35. Bolognese JA, Schnitzer TJ, Ehrich EW. Response relationship of VAS and Likert scales in osteoarthritis efficacy measurement. *Osteoarthritis Cartilage*. 2003;11(7):499–507.
36. Averbuch M, Katzper M. Assessment of visual analog versus categorical scale for measurement of osteoarthritis pain. *J Clin Pharmacol*. 2004;44(4):368–372.