The Journal of Rheumatology 2021;xx:xxxx
doi:10.3899/jrheum.200928
First Release April 15 2021

The Journal of Rheumatology

Comparing the Visual Analog Scale and the Numerical Rating Scale in Patient-reported Outcomes in Psoriatic Arthritis

Weiyu Ye[1], Simon Hackett[2], Claire Vandevelde[3], Sarah Twigg[4], Philip S Helliwell[5], and Laura C. Coates[2]

ABSTRACT. ***Objective.*** Patient self-report scales are invaluable in psoriatic arthritis (PsA), as they allow physicians to rapidly assess patient perspectives of disease activity. We aimed to assess the agreement of the visual analog scale (VAS), a 100-mm horizontal line, and the numerical rating scale (NRS), a 21-point scale ranging from 0 to 10 in increments of 0.5, in patients with PsA.

Methods. Data were collected prospectively across 3 UK hospital trusts from 2018 to 2019. All patients completed the VAS and NRS for pain, arthritis, skin psoriasis (PsO), and global disease activity. A subset completed an identical pack 1 week later. Demographic and clinical data were also collected. Agreement was assessed using medians and the Bland-Altman method. Intraclass correlation coefficients (ICCs) were used to assess test-retest reliability. Spearman rank correlation coefficients were used to assess dependency between scale scores and clinical variables.

Results. Two hundred ten patients completed the study; 1 withdrew consent. Thus, 209 were analyzed. For pain, arthritis, skin PsO, and global disease activity, the difference between the VAS and NRS lay mostly within 1.96 SD of the mean, suggesting reasonable agreement between the 2 scales. Among the patients, 64.1% preferred the NRS. The ICCs demonstrated excellent test-retest reliability for both VAS and NRS. Higher VAS and NRS scores were associated with increased tender/swollen joint count, poorer functional status, and greater life impact.

Conclusion. The VAS and NRS show reasonable agreement in key patient-reported outcomes in PsA. Results from both scales are correlated with disease severity and life impact.

Key Indexing Terms: numerical rating scale, patient-reported outcomes, psoriatic arthritis, self-assessment, visual analog scale

WY is a National Institute for Health Research (NIHR) Academic Clinical Fellow. LCC is an NIHR Clinician Scientist and Senior Clinical Research Fellow funded by an NIHR Clinician Scientist award. The research was supported by the NIHR Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. We acknowledge the support of the NIHR Clinical Research Network.

[1]W. Ye, NIHR Academic Clinical Fellow, MB BChir, Oxford University Clinical Academic Graduate School, John Radcliffe Hospital, Oxford; [2]S. Hackett, Academic Foundation Doctor, PhD, L.C. Coates, NIHR Clinician Scientist and Senior Clinical Research Fellow, PhD, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science, University of Oxford, The Botnar Research Centre, Oxford; [3]C. Vandevelde, Consultant Rheumatologist and Honorary Senior Lecturer, MD, NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, and Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds; [4]S. Twigg, Consultant Rheumatologist, MD, Bradford Teaching Hospitals NHS Foundation Trust, St. Lukes Hospital, Bradford; [5]P.S. Helliwell, Professor of Clinical Rheumatology, PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, Leeds, UK.

Address correspondence to Dr. W. Ye, Oxford University Clinical Academic Graduate School, Room 3A31, The Cairns Library IT Corridor, Level 3, John Radcliffe Hospital, Oxford, OX3 9DU, UK.
Email: christina.ye@conted.ox.ac.uk.

Accepted for publication November 10, 2020.

Psoriatic arthritis (PsA) is a chronic inflammatory arthritis with a significant effect on daily function and quality of life.[1,2,3] To evaluate this in clinical practice, patient self-report scales carry immense value. A commonly used validated scale is the visual analog scale (VAS), which consists of a 100-mm horizontal line that patients mark based on symptom severity.[4,5] This can be susceptible to random errors during completion or measurement, and systematic error, as photocopying can alter line length.[6]

The numerical rating scale (NRS) is a 21-point horizontal scale ranging from 0 to 10 in increments of 0.5, with higher numbers indicating greater severity.[7,8] Compared to the VAS, it is simpler to complete, faster to score, and less susceptible to measurement error.[9,10]

To date, the NRS has been validated for outcome measures in ankylosing spondylitis (AS).[9] Although the construct validity of an NRS of patient global health assessment has been validated in PsA,[11] the NRS and VAS have not yet been directly compared. Our study aimed to assess the agreement and reliability of the VAS and NRS in PsA for all key outcomes measured using the VAS, and to correlate the results with other clinical measurements and patient outcomes.

METHODS

Study design. We conducted a questionnaire study comparing the VAS to the

NRS for pain, arthritis, skin psoriasis (PsO), and global disease activity in patients aged ≥ 18 years with definite PsA (according to the ClASsification of Psoriatic ARthritis [CASPAR] criteria[12] or previous diagnosis by a rheumatologist). Patients were recruited from 3 UK hospital trusts (Oxford University Hospitals, Leeds Teaching Hospitals, and Bradford Teaching Hospitals) from December 20, 2018, to August 22, 2019.

All patients completed both scales in 1 clinic visit within usual care. The order the scales were presented alternated throughout the questionnaire. Patients were given an identical pack with a prepaid self-addressed envelope with instructions to return the completed questionnaires 1 week later. This ceased when returned questionnaire numbers were sufficient to evaluate test-retest reliability. The 1-week timepoint was chosen as it was assumed that most patients' disease activity state will have not changed significantly. This was clarified with an additional question on disease activity at 1 week.

Information was collected on patient demographics and treatment according to a standard protocol. Patients self-rated their disease severity as "unnoticeable," "very mild," "mild," "moderate," or "severe," and were asked to indicate their preferred scale. The effect of disease was assessed using the Psoriatic Arthritis Impact of Disease 12-item questionnaire (PsAID-12),[13] which has a validated, patient acceptable symptom state (PsAID-12 score ≤ 4) to stratify high-impact and low-impact disease. Functional status was assessed using the Health Assessment Questionnaire–Disability Index (HAQ-DI).[14] Patients were also examined by the treating rheumatologist for tender and swollen joint count (TJC/SJC), skin PsO body surface area (BSA), Leeds Enthesitis Index (LDI), and dactylitis count.

Statistical analyses. Median and IQR for all variables were calculated due to the nonparametric distribution of data. Variability between the VAS and NRS were assessed using the Bland-Altman method, which plots the mean of the scale scores against their difference.[15] The limits of agreement are defined as ± 1.96 SD of the mean. For reasonable agreement, points should lie within the limits approximately 95% of the time. Intraclass correlation coefficients (ICCs; 2-way mixed model absolute agreement) were used to assess test-retest reliability, with ICCs > 0.75 considered to demonstrate concordance.[9,16] Spearman rank was used to assess correlation between different variables. All analyses were performed using R (version 3.6.1; R Core Team).

Ethical considerations. This study was approved by the London-Surrey Research Ethics Committee (reference 18/LO/2057). All patients gave written informed consent.

RESULTS

Patients. Two hundred ten patients completed the clinic visit; 1 withdrew consent. Thus, data from 209 patients were analyzed. Of these, 60.0% were male, with a mean age of 51.7 years and a median PsA duration of 7.0 years. Separating by PsA subtype, 84.7% had peripheral, 9.1% had axial, 1.4% had enthesitis predominant, and 4.3% had ≥ 2 subtypes. Further, 88.5% of patients had limited/no skin PsO, 9.6% had extensive skin PsO (6–20% BSA), and 1.4% had very extensive skin PsO (> 20% BSA). In terms of treatment, 17.7% of patients were treated with nonsteroidal antiinflammatory drugs, 54.5% with conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), 49.3% with biologics, and 21.5% with combination csDMARD plus biologics.

Sixty-two of 107 patients given an identical pack to complete at 1 week returned the questionnaires. Of these, 61.2% were male, with a mean age of 57.2 years and a median PsA duration of 10 years. Their clinical and treatment characteristics were broadly representative of the whole cohort. Thirty-six patients responded to the question assessing their current disease activity compared to their clinic visit (stable = 26, improvement = 3, deterioration = 7).

Scale scores and agreement. The median VAS and NRS for all variables are detailed in Table 1. Median NRS scores tend to be slightly higher than VAS scores. The variability appears to be greatest for global disease activity, and least for skin PsO (Figure 1). For all 4 variables, there appears to be reasonable agreement between the 2 scales.

In clinic, 64.1% patients preferred the NRS over the VAS. At 1 week, although a greater proportion preferred the NRS, 14 of 62 patients had changed their preference.

Test-retest reliability. Comparing clinic and 1-week VAS scores, the ICCs for pain, skin PsO, arthritis, and global disease activity were 0.91 (95% CI 0.84–0.94), 0.93 (95% CI 0.87–0.96), 0.85 (95% CI 0.74–0.91), 0.89 (95% CI 0.81–0.93), respectively. For NRS, the ICCs for pain, skin PsO, arthritis, and global disease activity were 0.93 (95% CI 0.88–0.96), 0.89 (95% CI 0.82–0.94), 0.91 (95% CI 0.84–0.94), 0.91 (95% CI 0.85–0.95), respectively (data not shown). This suggests both scales have excellent test-retest reliability. The results were similar in patients who reported no change in disease activity at 1 week (n = 26, VAS and NRS ICCs for all variables were ≥ 0.92).

Correlation with disease activity and other clinical outcomes. Table 2 details the median VAS and NRS scores in clinic for all variables, separated by patients' self-reported disease activity. Compared to the Unnoticeable group, patients in the Severe group had higher VAS and NRS scores. They also had increased TJC/SJC and higher PsAID-12 and HAQ-DI scores, suggestive of greater life impact and poorer functional status.

Using Spearman rank, we found statistically significant correlations between clinic global disease activity VAS scores and TJC (r_s 0.55, $P < 0.001$), SJC (0.49, $P < 0.001$), tender entheseal points (0.40, $P < 0.001$), HAQ-DI score (0.64, $P < 0.001$), and PsAID-12 scores (0.87, $P < 0.001$). Similar trends were found between clinic global disease activity NRS scores and TJC (0.52, $P < 0.001$), SJC (0.44, $P < 0.001$), tender entheseal points (0.43, $P < 0.001$), HAQ-DI score (0.68, $P < 0.001$), and PsAID-12 scores (0.90, $P < 0.001$; data not shown). Notably, correlation between VAS and NRS scores with dactylitis count was not statistically significant ($P > 0.05$). This may be due to

Table 1. Median VAS and NRS scores in clinic and home 1 week later.

	Clinic, n = 209		Home, n = 62	
Variable, median (IQR)	VAS	NRS	VAS	NRS
Pain	3.0 (5.9)	3.0 (5.5)	3.5 (5.7)	4.0 (6.0)
Skin PsO	1.9 (5.4)	2.5 (6.5)	2.9 (6.1)	3.0 (5.5)
Arthritis	3.2 (6.0)	3.3 (6.0)	3.2 (5.4)	3.8 (5.4)
Global disease activity	3.1 (5.7)	4.0 (6.5)	3.1 (5.3)	4.0 (5.5)
Preferred scale, n (%)	64 (30.6)	134 (64.1)	26 (41.9)	31 (50.0)

NRS: numerical rating scale (score 0–10); PsO: psoriasis; VAS: visual analog scale (score 0–10).

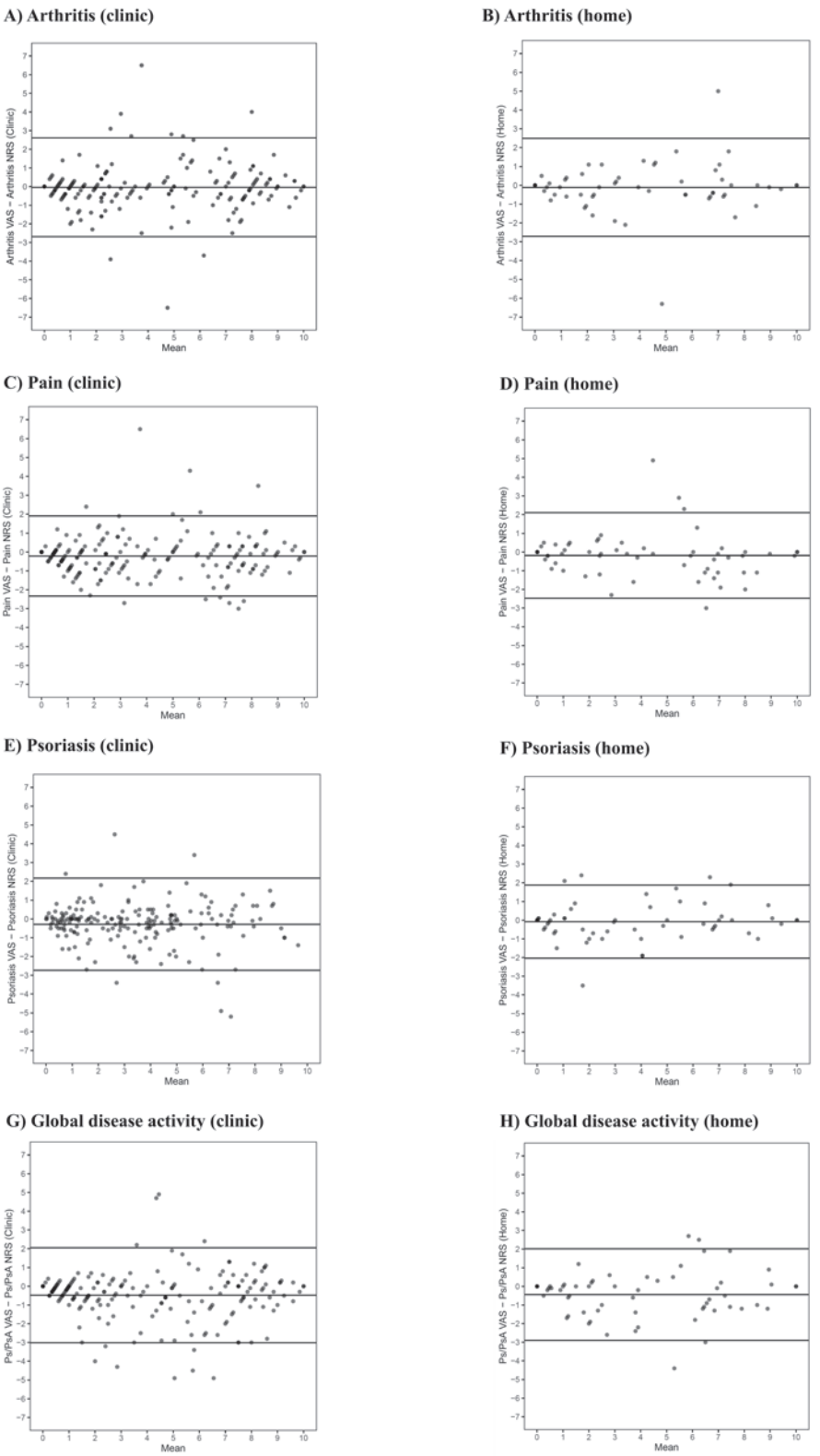

Figure 1. Bland-Altman plots comparing VAS and NRS in clinic and at home 1 week later. VAS: visual analog scale; NRS: numerical rating scale.

active dactylitis being uncommon in our cohort, with only 10 patients experiencing active dactylitis. There was also no statistically significant correlation between VAS and NRS scores with age or disease duration (all $P > 0.05$). Collectively, these results suggest that results from both VAS and NRS may be taken as a crude correlate of disease activity.

Table 2. Clinical assessment outcomes according to patients' self-reported disease activity.

Variable, Median (IQR)	Unnoticeable	Very Mild	Mild	Moderate	Severe
Pain VAS (0–10)	0.0 (0.5)	0.7 (1.1)	2.8 (3.9)	5.3 (3.4)	8.7 (1.7)
Pain NRS (0–10)	0.3 (0.5)	1.5 (1.5)	3.0 (3.0)	5.0 (3.5)	8.8 (1.1)
Psoriasis VAS (0–10)	0.3 (0.6)	0.8 (1.3)	2.4 (4.5)	4.7 (5.7)	8.7 (1.4)
Psoriasis NRS (0–10)	0.3 (0.5)	1.0 (1.5)	3.0 (3.6)	5.0 (6.4)	8.8 (2.5)
Arthritis VAS (0–10)	0.1 (0.5)	1.0 (1.3)	2.9 (4.6)	6.2 (3.1)	8.8 (1.9)
Arthritis NRS (0–10)	0.0 (0.6)	1.0 (1.5)	3.5 (4.0)	6.0 (3.5)	8.5 (1.1)
Global disease activity VAS (0–10)	0.1 (0.3)	0.9 (1.3)	3.1 (4.2)	5.4 (3.6)	8.7 (1.1)
Global disease activity NRS (0–10)	0.5 (0.6)	1.0 (1.0)	3.5 (4.0)	6.5 (3.0)	9.0 (1.1)
Tender joint count (0–68)	0.0 (0.0)	0.0 (1.5)	0.0 (2.0)	2.0 (4.0)	6.5 (10.5)
Swollen joint count (0–66)	0.0 (0.0)	0.0 (1.0)	0.0 (2.0)	0.5 (3.0)	2.5 (5.0)
Dactylitis count (0–20)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Tender entheseal points (0–6)	0.0 (0.0)	0.0 (0.0)	0.0 (1.0)	0.0 (1.3)	0.0 (1.0)
PsAID-12 score (0–10)	0.3 (0.6)	1.6 (1.7)	3.0 (3.0)	5.1 (3.3)	7.9 (2.4)
HAQ-DI score (0–3)	0.0 (0.1)	0.2 (0.8)	0.5 (0.8)	0.9 (0.9)	2.1 (0.9)

HAQ: Health Assessment Questionnaire–Disability Index; NRS: numerical rating scale; PsAID-12: Psoriatic Arthritis Impact of Disease 12-item questionnaire; VAS: visual analog scale.

DISCUSSION

To our knowledge, our study is the first to compare the VAS and NRS in patients with PsA. We demonstrate that both scales show high levels of agreement in patient-reported pain, skin PsO, arthritis, and global disease activity, and that both have excellent test-retest reliability. Overall, patients indicated a preference for the NRS over the VAS.

Our findings are consistent with previous studies. Price, *et al*[17] demonstrated that the NRS and VAS are correlated in the measurement of pain in patients with orofacial pain. Van Tubergen, *et al*[9] found that the NRS and VAS of the Bath AS Disease Activity Index, Bath AS Functional Index, and Dougados Functional Index showed high levels of agreement in 536 patients with AS. They also found that patients preferred the NRS.

As expected, patients' self-reported disease severity matched the severity of VAS and NRS scores. Higher scores and self-reported disease severity were also associated with greater clinical correlates of active inflammation, and correspondingly, poorer daily function and greater life impact. Those in the self-reported "moderate" and "severe" disease severity categories had median HAQ-DI scores of 0.9 and 2.1, respectively, with median PsAID-12 scores of 5.1 and 7.9, respectively. These results echo those from a recent Singaporean study,[11] which observed that the NRS of a patient global assessment was strongly correlated with physical and mental function, as assessed by the Medical Outcomes Study 36-item Short Form Health Survey. It can also differentiate between different levels of disease severity, defined using composite scores including the HAQ-DI, Disease Activity Score in 28 joints, and minimal disease activity criteria.

Strengths of our study include recruitment of patients from 3 separate centers and comparing the scales in an unselected group of patients with PsA within routine clinical practice. Limitations include the sample size being too small to enable detailed analyses of more uncommon features, such as active dactylitis, lack of assessment of sensitivity to change, and the absence of objective measures of inflammation such as blood results and imaging. Moreover, the 1-week interval between questionnaires also meant some patients felt their disease activity state had changed.

In conclusion, our results suggest that the NRS and VAS are comparable. This is relevant to both clinical and research settings, where scales are routinely utilized to assess patients' perspectives of disease activity and to evaluate treatment effect.

REFERENCES

1. Kwok T, Pope JE. Minimally important difference for patient-reported outcomes in psoriatic arthritis: Health Assessment Questionnaire and pain, fatigue, and global visual analog scales. J Rheumatol 2010;37:1024-8.
2. Coates LC, Helliwell PS. Psoriatic arthritis: state of the art review. Clin Med 2017;17:65-70.
3. Ritchlin C, Scher JU. Strategies to improve outcomes in psoriatic arthritis. Curr Rheumatol Rep 2019;21:72.
4. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17:45-56.
5. Huskisson EC, Jones J, Scott PJ. Application of visual-analogue scales to the measurement of functional capacity. Rheumatol Rehabil 1976;15:185-7.
6. Dixon JS, Bird HA. Reproducibility along a 10 cm vertical visual analogue scale. Ann Rheum Dis 1981;40:87-89.
7. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain. 1986;27:117-26.
8. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? Pain 1994; 58:387-92.
9. Van Tubergen A, Debats I, Ryser L, Londoño J, Burgos-Vargas R, Cardiel MH, et al. Use of a numerical rating scale as an answer modality in ankylosing spondylitis-specific questionnaires. Arthritis Rheum 2002;47:242-8.
10. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al; European Palliative Care Research Collaborative (EPCRC). 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.

11. Leung YY, Ho KW, Zhu TY, Tam LS, Kun EW, Li EK. Construct validity of the modified numeric rating scale of patient global assessment in psoriatic arthritis. J Rheumatol 2012;39:844-848.
12. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H; CASPAR Study Group. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665-73.
13. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, et al; EULAR PsAID Taskforce. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. Ann Rheum Dis 2014;73:1012-9.
14. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307-10.
16. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-63.
17. Price DD, Bush FM, Long S, Harkins SW. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. Pain 1994;56:217-26.