Suitability Of The Numerical Pain Rating Scale For Measuring Pain in Clinical Trials Evaluating Interventions For People With Shoulder Disorders According To The OMERACT Filter 2.2

Hana Marmura[1], Rochelle Furtado[2] Romi Haas[3], Sofia Ramiro[4], Joel Gagnier[5], Arianne Verhagen[6], Samuel Whittle[3,7], Dorcas Beaton[8], Beverley Shea[9], Pamela Richards[10], Rachelle Buchbinder[3]
[1]University of North Carolina at Chapel Hill, NC, United States, [2]University Health Network, Toronto, ON, Canada, [3]Monash University, Melbourne, Australia. [4]Leiden University Medical Center, Heerlen, The Netherlands, [5]Western University, London, ON, Canada, [6]University of Technology Sydney, Sydney, Australia, [7]The Queen Elizabeth Hospital, Adelaide, Australia, [8]Institute of Health & Work and the University of Toronto, Toronto, ON, Canada, [9]Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada, [10]Patient research partner
hmarmura@unc.edu

Disclosures: Rochelle Furtado (N), Hana Marmura (N), Romi Haas (N), Sofia Ramiro (AbbVie Inc, Johnson & Johnson MedTech, Merck & co, Novartis Pharmaceuticals Corporation, Alfasigma SpA, Pfizer Inc., UCB Pharma SA. Assessment of Spondyloarthritis International Society and EULAR Quality of Care Committees), Joel Gagnier (N), Arianne Verhagen (N) Samuel Whittle (N), Dorcas Beaton (OMERACT board member), Beverley Shea (N), Pamela Richards (N), Rachelle Buchbinder (N).

INTRODUCTION: The Outcome Measures in Rheumatology (OMERACT) group recommends pain as a core domain in clinical trials for shoulder conditions. This study evaluated the suitability of the Numerical Pain Rating Scale (NPRS) for measuring pain using the OMERACT Filter 2.2.

METHODS: Following the OMERACT Handbook for instrument selection, a systematic review assessed the NPRS's construct validity, reliability, longitudinal construct validity, ability to discriminate in clinical trials, and thresholds of meaning for shoulder disorders. Articles were independently screened, appraised with the COSMIN-OMERACT Good Methods Checklist**,** and rated as green (good), amber (caution), red (poor), or white (no evidence). Findings were summarized in a Summary of Measurement Properties table and discussed at the 2025 OMERACT Shoulder Special Interest Group.

RESULTS SECTION: The NPRS was rated amber for domain match and green for feasibility. Twelve studies met eligibility criteria for the systematic review. Three studies examined construct validity, two test-retest reliability, four longitudinal construct validity, six clinical trial discrimination and three thresholds of meaning. Only 6/18 components of evidence had good methods, with a further 12 components rated amber (Table 1). NPRS versions and formats varied substantially across studies. Consequently, 75% of respondents at the 2025 OMERACT conference agreed that synthesizing results was inappropriate due to this heterogeneity.

DISCUSSION: The NPRS cannot progress to the endorsement stage of instrument selection for assessing pain for shoulder conditions due to heterogeneity among NPRS versions used within primary studies and insufficient evidence for any single NPRS version. Each version should be considered a distinct instrument, and current evidence is insufficient to support its use in clinical trials.

SIGNIFICANCE: This study has important implications for next steps in instrument selection for the pain domain in a core outcome for clinical trials assessing those with shoulder disorders. Other single-item instruments measuring pain may have similar limitations as the NPRS given inconsistent format and reporting of these scales. Therefore, we will focus on assessing the suitability of multi-item pain scales such as the subscale of the Shoulder Pain and Disability Index and the Shoulder Pain Score as they may be reported more consistently across literature.

Table 1. Summary of Measurement Properties Table for the Numerical Pain Rating Scale.

Author	Year	NPRS Brief Description	NPRS Pain Category	Construct Validity	Inter-method reliability	Test- retest reliability	Longitudinal construct validity	Clinical trial discrimination	Thresholds of meaning
Observational Studies									
Hummel-Berry	2001	Pain at rest, Pain during motion	Background Evoked	±					
vanMeeteren	2006	Pain on movement	Evoked				+		
Mintken	2009	Not described	NS			±	-	-	+
Michener	2011	Pain at rest, Pain with ADLs, Pain with strenuous activity, NPRS average of above 3 items	Averaged (only averaged pain reported/used in analyses)				+	+	+
Secondary Analyses (of Cohort Studies or Randomized Controlled Trials)									
Tubach	2006	Pain on movement	Evoked						±
O'Halloran	2013	Not described	NS				±		
Riley	2015	Average pain in past 24h, Present pain at rest, Pain with shoulder elevation, Hawkins-Kennedy test, and Neer test	Averaged Background Evoked	-					
Riley	2019	Pain at time of testing	Averaged	±		+			
Randomized Controlled Trials									
Atkinson	2008	Not described	NS					-	
Rhon	2014	Not described	NS					+	
Kardouni	2015	Current pain (different wording pre and post intervention)	Averaged					+	
Ranagan	2020	Not described	NS					±	
Synthesis									
Studies per property (n)				3	0	2	4	6	3
Synthesis statement				Synthesis per property was not possible due to heterogeneity in NPRS instruments.					