

Advancing Stiffness Measurement in Rheumatic Disease: Report from the Stiffness Special Interest Group at OMERACT 2018

Ethan T. Craig, Ana-Maria Orbai , Sarah Mackie, Susan J. Bartlett, Clifton O. Bingham III, Susan Goodman, Catherine Hill, Robert Holt, Amye Leong, Chetan Karyekar, Ying Ying Leung , Pamela Richards, and Serena Halls

ABSTRACT. Objective. To improve measurement of stiffness in rheumatic disease.

Methods. Data presented included (1) 2 qualitative projects, (2) the rheumatoid arthritis (RA) stiffness patient-reported outcome measure (RAST), and (3) 3 items assessing stiffness severity, duration, and interference.

Results. Stiffness is multidimensional and includes aspects of stiffness experience such as duration, severity, and effect. Stiffness items showed construct validity in RA. Further efforts are required to develop an instrument that will be taken through the Outcome Measures in Rheumatology (OMERACT) Filter 2.1 for instrument selection.

Conclusion. The research agenda for the group includes domain content voting for individual diseases, and development of stiffness item banks and disease-specific short forms. (J Rheumatol First Release May 1 2019; doi:10.3899/jrheum.181074)

Key Indexing Terms:

OUTCOME ASSESSMENT

MORNING STIFFNESS

OMERACT

From the Johns Hopkins University School of Medicine, Baltimore, Maryland; Department of Rheumatology, Hospital for Special Surgery, Weill Cornell Medical College, New York, New York; Horizon Pharma Inc., Lake Forest; College of Pharmacy, University of Illinois, Chicago, Illinois; Healthy Motivation, Bone and Joint Decade Global Alliance for Musculoskeletal Health, Santa Barbara, California; Janssen Scientific Affairs LLC, Horsham, Pennsylvania, USA; Leeds Institute of Rheumatology and Musculoskeletal Medicine, University of Leeds; and Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds; University of Bristol; University of the West of England, Bristol, Bristol, UK; McGill University, Montreal, Quebec, Canada; Division of Medicine, The University of Adelaide, Adelaide; Rheumatology Unit, The Queen Elizabeth Hospital, Woodville, Australia; Singapore General Hospital, Duke-National University of Singapore (NUS) Medical School, Singapore.

AMO is a Jerome L. Greene Foundation Scholar and is supported in part by a research grant from the US National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health (NIH) under award number P30-AR070254 (Core B), a Rheumatology Research Foundation Scientist Development award, and a Staurulakis Family Discovery award. Work from the US Cohort was supported by the NIAMS of the NIH under award numbers [P30-AR070254 Core B] and [P30-AR053503 Core D] and Patient Centered Outcomes Research Institute (PCORI) Pilot Project Award number [IP2-PI000737], and the Camille Julia Morgan Arthritis Research and Education Fund. All statements in this report including its conclusions are the opinions of the authors and do not necessarily reflect those of PCORI, its board of governors, or its methodology committee, or of NIH or NIAMS. RH is an employee of Horizon Pharma LLC. CK is an employee of Janssen Scientific Affairs LLC.

E.T. Craig, MD, MHS, Johns Hopkins University School of Medicine; A.M. Orbai, MD, MHS, Johns Hopkins University School of Medicine; S. Mackie, PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust; S.J. Bartlett, PhD, McGill University, and Johns Hopkins University School of Medicine; C.O. Bingham III, MD, Johns Hopkins University School of Medicine; S. Goodman, MD, Department of Rheumatology, Hospital for Special

Surgery, Weill Cornell Medical College; C. Hill, MBBS, MD, MSc, FRACP, Division of Medicine, The University of Adelaide, and Rheumatology Unit, The Queen Elizabeth Hospital; R. Holt, PharmD, MBA, Horizon Pharma Inc., and Adjunct Professor, College of Pharmacy, University of Illinois; A. Leong, MBA, Healthy Motivation, Bone and Joint Decade Global Alliance for Musculoskeletal Health; C. Karyekar, MD, PhD, Janssen Scientific Affairs LLC; Y.Y. Leung, MD, MB ChB, MRCP, FHKAM, FAMS, Singapore General Hospital, Duke-NUS Medical School; P. Richards, Patient Research Partner, University of Bristol; S. Halls, PhD, University of the West of England. Dr. Craig and Dr. Orbai are co-first authors.

Address correspondence to Dr. E.T. Craig, Medicine Business Office (111), 3900 Woodland Ave., Philadelphia, Pennsylvania 19104, USA. E-mail: ethan.craig@uphs.upenn.edu

Accepted for publication January 30, 2019.

In 2016, the Outcome Measures in Rheumatology (OMERACT) Stiffness Special Interest Group (SIG) was formed to develop an appropriate instrument to measure stiffness in rheumatoid arthritis (RA)¹. The SIG reviewed results from a systematic literature review of stiffness patient-reported outcome measures (PROM), and noted a dearth of data regarding content validity, construct validity, responsiveness, or reliability of any of these instruments². Thus, the available measures did not pass the OMERACT Filter 2.0 “eyeball test” for good match with the domain of interest. A longterm research agenda was proposed including development of stiffness content mapping through qualitative research followed by development and validation of stiffness assessment tools in RA.

Stiffness is a common and important symptom in

rheumatic diseases, including RA, spondyloarthritis, and polymyalgia rheumatica (PMR)^{3,4,5}, and is present even in low disease activity states⁶. Stiffness was initially included in the 1987 American College of Rheumatology classification criteria for RA and is included in the Bath Ankylosing Spondylitis Disease Activity Index, RA Flare Questionnaire, and OMERACT core domain sets for RA flare, ankylosing spondylitis (AS), psoriatic arthritis, and PMR^{7,8,9,10,11}. Stiffness has been identified by patients as a hallmark of the RA experience, and is associated with flares, decreased physical function, and effect upon daily functioning¹². Prospective observational studies and survey data have demonstrated that stiffness duration is independently associated with decreased work productivity and early retirement^{13,14,15}.

Considerable qualitative work has been performed by our working group to improve our understanding of the patient experience of stiffness. Independent qualitative work in the United Kingdom and United States identified that in RA, stiffness is a variable and complex symptom that affects patients' daily lives and is not exclusive to the morning hours^{6,12}. Conceptual models developed in these studies have been discussed, synthesized, and expanded on across rheumatic conditions^{1,2}.

Despite its central involvement in patient experience, measurement of stiffness remains suboptimal in clinical trials, and no validated instruments exist for its measurement.

At the 2018 OMERACT Stiffness SIG, 37 participants were present, including 22 healthcare providers, 7 fellows, 4 industry representatives, and 4 patient research partners (PRP). The primary aims of the OMERACT 2018 Stiffness SIG were to (1) gain input regarding comprehensive understanding of the domain of stiffness, (2) review data regarding preliminary stiffness scales from 2 cohorts of patients with RA, and (3) obtain feedback on other areas of development, including contextual factors affecting stiffness.

Furthering Understanding of the Domain of Stiffness

The dimensions (or subdomains) of the experience of stiffness that are relevant for stiffness measurement have yet to be endorsed by the OMERACT community. Stiffness occurs in many disease states, but patient experience among these diseases may differ. There remains, therefore, a fundamental question of whether stiffness is optimally measured using disease-specific or universal measures. The first task will be to identify themes of commonality and divergence among diseases. To accomplish this, we will reconsider qualitative data to identify themes that were repeated across diseases, and those that were unique to specific diseases. These themes will be used to generate items, and a Delphi process will be undertaken with patients with each disease of interest. The Delphi process will aim to identify items that achieve consensus among patients from varied disease states and those that achieve consensus only among patients with

individual diseases. Figure 1 summarizes the proposed domain generation and Delphi process.

Review of Measurement Properties of Stiffness

Instruments

Data from 2 instruments were considered. These instruments have been developed based on qualitative studies of stiffness in RA and consist of the RA stiffness patient-reported outcome (PRO) measure (RAST) and 3 stiffness numerical rating scale (NRS) items. Their measurement properties as presented during the SIG are briefly described below.

Measurement properties of the RAST. Work in the United Kingdom has focused on the development of an instrument guided by the patient experience. This study was approved by the University of the West of England, Bristol IRB, approval # 14/WA/1162. Written informed consent was obtained from all patients.

This body of work has included qualitative investigation into the patient perspective of stiffness⁶, followed by patient focus groups, an iterative process of item development involving clinicians, researchers, and patients, and subsequent cognitive interviews to refine draft items¹⁶. A postal survey and analysis to develop the new instrument called "RAST" was presented at OMERACT 2016. This 21-item questionnaire addresses 3 components of stiffness: severity, physical effect, and psychosocial effect.

A key element of outcome measure development and testing is ensuring that outcome measures have appropriate measurement properties. One approach to ensure that outcome measures are developed appropriately and have suitable measurement properties is the OMERACT Filter 2.1. The OMERACT Filter recommends that outcome measures are evaluated against the concepts of truth (is the measure identifying what is intended in an unbiased manner?), discrimination (is the measure sensitive and reliable?), and feasibility (is the measure understandable and time efficient for use in clinical and research environments?)^{17,18}. Work presented at the OMERACT 2018 meeting focused on the match between the RAST and the OMERACT Filter 2.1 to identify the measurement properties that are currently met and those that require further development. Figure 2^{2,6,12} demonstrates the match between the OMERACT Filter 2.0 instrument selection algorithm and the RAST evidence from the presented development study. The RAST data provided new evidence for the concepts of truth and feasibility. Data supporting discrimination are the target of future work.

Measurement properties of 3 stiffness items. A complementary approach was used in the US cohort to further analyze the domain of stiffness. Three stiffness items using a 5-point NRS, assessing severity, duration, and interference with daily activities, were administered in a cohort of patients with RA. These items were developed based upon the same conceptual model and upon qualitative data¹². A 100-mm

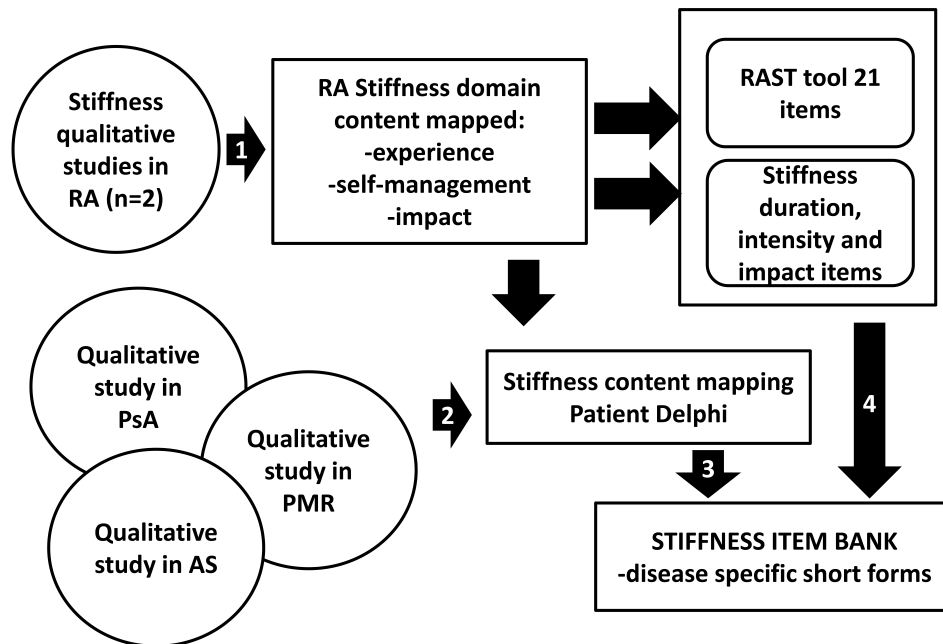


Figure 1. Stiffness in rheumatic diseases study design. Step 1 designates the qualitative analysis step that led to determining the stiffness domain content specifically in RA and 2 draft instruments for their use. Step 2 illustrates qualitative studies existing in other diseases that we will use to define stiffness content in other rheumatic diseases, followed by a patient Delphi. We will then develop disease-specific items and together with the existing RA stiffness items we will create a stiffness item bank. Short forms from this item bank can then be individually assessed through OFISA. RA: rheumatoid arthritis; RAST: RA stiffness patient-reported outcome measure; PsA: psoriatic arthritis; AS: ankylosing spondylitis; PMR: polymyalgia rheumatica; OFISA: OMERACT Filter 2.1 Instrument Selection Algorithm; OMERACT: Outcome Measures in Rheumatology.

visual analog scale (VAS) rating stiffness intensity was subsequently added and completed by a subset of patients.

1. In the past 7 days, how long did your stiffness last on average? (0: none, 1: < 30 min, 2: 30 min to 1 h, 4: > 1–2 h, 5: > 2–4 h, 6: > 4 h)

2. In the past 7 days, how would you rate your stiffness on average? (0: none, 1: mild, 2: moderate, 3: severe, 4: very severe)

3. In the past 7 days, how often did your stiffness interfere with your activities? (0: never, 1: rarely, 2: sometimes, 3: often, 4: always)

4. How would you describe your stiffness in the past week? (VAS, 0 none to 100 extreme)

Items were administered as part of an observational study evaluating the use of PRO in routine care at an academic rheumatology clinic, along with selected Patient-Reported Outcome Measurement Information System (PROMIS) measures [pain interference, fatigue, physical function, sleep disturbance, depression, anxiety (all version 1.0), and ability to participate in social roles (v2.0)], a 100-mm pain VAS, and a health transition item^{19,20}. Our study was approved by the Johns Hopkins IRB, approval # NA_00040493. Written informed consent was obtained from all patients. In total, 196 patients with RA completed the baseline visit and 83 participants completed both VAS and NRS items.

Data offered preliminary evidence of construct validity based upon correlation with other PROM expected to correlate with stiffness. Responsiveness in states of change in several clinical anchors and longitudinal construct validity was also seen for all 3 items.

Among the 83 participants who completed both VAS and NRS items, the stiffness VAS, which did not specify a subdomain, correlated most strongly with the stiffness severity score (Figure 3). A score of “none” in severity corresponded to a median VAS of 0, “mild” to a VAS of 15, “moderate” to a VAS of 60, and “severe” to a VAS of 85.

The data from the US cohort lend additional support toward a multidimensional approach to the measurement of stiffness.

Further SIG Discussion

PRP emphasized 2 major trends. First, there was considerable frustration over the ongoing use of morning stiffness duration in the literature and underscored findings from qualitative work that stiffness fluctuated throughout the day and was not restricted to the morning. Second, there is a need to better understand contextual factors related to stiffness. For example, one PRP pointed out the involvement of medications in modifying stiffness. Another PRP noted the varied experience of stiffness depending upon the context of the current level of functioning (“It really depends on what I was doing the week

RAST OMERACT SUMMARY OF MEASUREMENT PROPERTIES							
Author/year	Truth Domain match	Feasibility	Truth	Discrimination			
			Construct validity	Test retest reliability	Longitudinal construct validity	Clinical trial discrimination	Thresholds of meaning
Halls S, et al 2014	+						
Orbai AM, et al 2014	+						
Halls S, et al 2017			+				
Total available studies for each property	4	0	1	2	0	0	0
Total studies available for synthesis	4	0	1	2	0	0	0
Rating	GREEN	WHITE	AMBER	WHITE	WHITE	WHITE	WHITE

Figure 2. OMERACT summary of evidence for measurement properties of rheumatoid arthritis stiffness (RAST) questionnaire. Color designates quality of evidence: green = good methods used, use this evidence; amber = some cautions but we will use this evidence. In the rating row, color designates overall evidence-based instrument rating for the core instrument set: green = at least 2 pieces of evidence with good methods and consistent findings of adequate or better performance; amber = in-between green and red; red = inadequate performance on at least 1 study that used good methods; white = no evidence. Arithmetic signs designate the performance of the instrument according to that study (for each measurement property studied): + = adequate or better performance; ± = equivocal performance; - = less than adequate/poor performance.

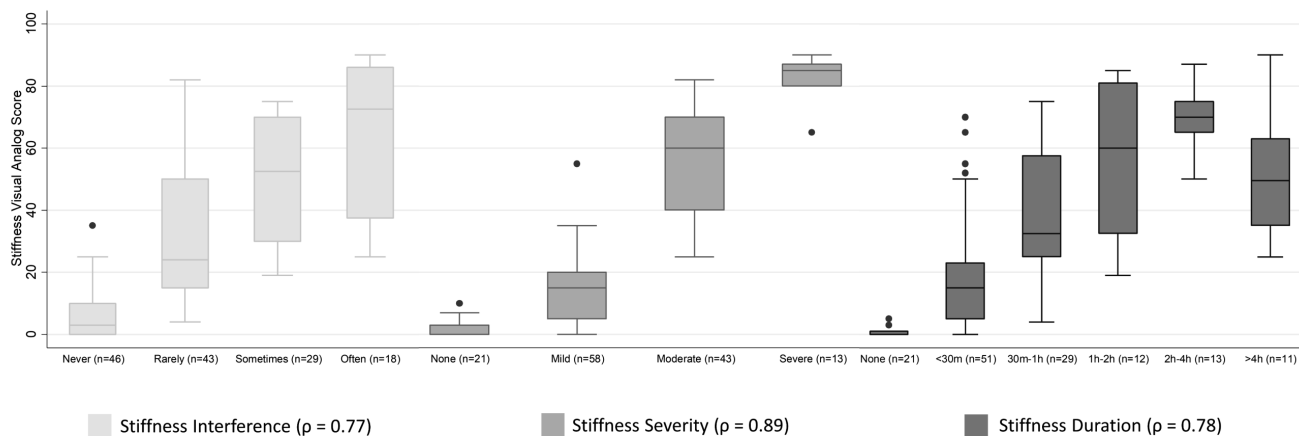


Figure 3. Correlation of stiffness VAS with stiffness NRS items. All correlations are Spearman ρ . VAS: visual analog scale; NRS: numerical rating scale.

before. Was I doing Yoga or in bed in a massive flare?"). Future efforts will also focus upon identifying contextual factors that modify the experience of stiffness.

Research Agenda

The OMERACT 2018 Stiffness SIG included the following items on its research agenda:

1. Re-evaluate qualitative data to delineate common and distinct domains (or subdomains) of the experience of stiffness across disease states.
2. Through a Delphi process, gain endorsement from the OMERACT community regarding the essential domains (or subdomains) for the measurement of stiffness.

3. Further investigate the 2 proposed stiffness instruments including a longitudinal survey to establish reliability and responsiveness of the RAST.

DISCUSSION

The Stiffness Special Interest Group will delineate the minimum necessary stiffness domain content for RA and additional rheumatic diseases. Once domain content has been defined, we will look for existing instruments to fulfill this need or likely develop a stiffness measurement instrument and build a body of evidence to support it meeting the OMERACT Filter 2.1 framework for instrument selection.

ACKNOWLEDGMENT

The authors acknowledge all OMERACT 2016 Stiffness Special Interest Group participants for their contributions. A travel scholarship from Janssen supported the fellow attendance to the OMERACT conference. A research grant from Horizon pharma provided funding in part for statistical analysis. All analysis was conducted by investigators. We thank Jennifer Horonjeff, PhD, for her collaboration with the SIG.

REFERENCES

- Orbai AM, Halls S, Hewlett S, Bartlett S, Leong A, Bingham III C; RA Flare Group Steering Committee. More than just minutes of stiffness in the morning: report from the OMERACT Rheumatoid Arthritis Flare Group stiffness breakout sessions. *J Rheumatol* 2015;42:2182-4.
- Halls S, Sinnathurai P, Hewlett S, Mackie SL, March L, Bartlett SJ, et al. Stiffness is the cardinal symptom of inflammatory musculoskeletal diseases, yet still variably measured: report from the OMERACT 2016 stiffness special interest group. *J Rheumatol* 2017;44:1904-10.
- Mackie SL, Arat S, da Silva J, Duarte C, Halliday S, Hughes R, et al. Polymyalgia rheumatica (PMR) special interest group at OMERACT 11: outcomes of importance for patients with PMR. *J Rheumatol* 2014;41:819-23.
- Calin A. The individual with ankylosing spondylitis: defining disease status and the impact of the illness. *Br J Rheumatol* 1995;34:663-72.
- van Tuyl LH, Lems WF, Boers M. Measurement of stiffness in patients with rheumatoid arthritis in low disease activity or remission: a systematic review. *BMC Musculoskelet Disord* 2014;15:28.
- Halls S, Dures E, Kirwan J, Pollock J, Baker G, Edmunds A, et al. Stiffness is more than just duration and severity: a qualitative exploration in people with rheumatoid arthritis. *Rheumatol* 2014;54:615-22.
- Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.
- van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. *J Rheumatol* 1997;24:2225-9.
- Orbai AM, de Wit M, Mease PJ, Callis Duffin K, Elmamoun M, Tillett W, et al. Updating the psoriatic arthritis (PsA) core domain set: a report from the PsA workshop at OMERACT 2016. *J Rheumatol* 2017;44:1522-8.
- Mackie SL, Twohig H, Neill LM, Harrison E, Shea B, Black RJ, et al; OMERACT PMR Working Group. The OMERACT core domain set for outcome measures for clinical trials in polymyalgia rheumatica. *J Rheumatol* 2017;44:1515-21.
- Bartlett SJ, Barbic SP, Bykerk VP, Choy EH, Alten R, Christensen R, et al. Content and construct validity, reliability, and responsiveness of the rheumatoid arthritis flare questionnaire: OMERACT 2016 workshop report. *J Rheumatol* 2017;44:1536-43.
- Orbai AM, Smith KC, Bartlett SJ, De Leon E, Bingham III CO. "Stiffness has different meanings, I think, to everyone": examining stiffness from the perspective of people living with rheumatoid arthritis. *Arthritis Care Res* 2014;66:1662-72.
- da Silva JA, Phillips S, Buttgerit F. Impact of impaired morning function on the lives and well-being of patients with rheumatoid arthritis. *Scand J Rheumatol Suppl* 2011;125:6-11.
- Mattila K, Buttgerit F, Tuominen R. Influence of rheumatoid arthritis-related morning stiffness on productivity at work: results from a survey in 11 European countries. *Rheumatol Int* 2015;35:1791-7.
- Westhoff G, Buttgerit F, Gromnica-Ihle E, Zink A. Morning stiffness and its influence on early retirement in patients with recent onset rheumatoid arthritis. *Rheumatology* 2008;47:980-4.
- Halls S, Dures E, Kirwan JR, Pollock J, Baker G, Edmunds A, et al. Development and testing of candidate items for inclusion in a new rheumatoid arthritis stiffness patient-reported outcome measure. *Rheumatology* 2018;57:263-72.
- Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO III, Conaghan PG, et al. The OMERACT Handbook. [Internet. Accessed May 17, 2017.] Available from: <https://omeract.org/resources>
- Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* 2014;67:745-53.
- Bingham CO 3rd, Bartlett SJ, Merkel PA, Mielenz TJ, Pilkonis PA, Edmundson L, et al. Using patient-reported outcomes and PROMIS in research and clinical applications: experiences from the PCORI pilot projects. *Qual Life Res* 2016;25:2109-16.
- Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS One* 2015;10:e0138543.