RHEUMATOLOGY

Editorial

doi:10.1093/rheumatology/kex247

Towards improving measurement of stiffness in rheumatology

This editorial refers to Development and testing of candidate items for inclusion in a new rheumatoid arthritis stiffness patient-reported outcome measure, Serena Halls *et al.* doi: 10.1093/rheumatology/kex085.

Stiffness is an important symptom in RA, endorsed by patients and physicians as a disease activity indicator [1] with associated life impact [2, 3]. Stiffness is important across rheumatic diseases, and its assessment informs clinical judgement about inflammatory activity. Stiffness, framed as morning stiffness duration, is a criterion in the 2012 ACR classification for PMR, and was historically a criterion in the RA classification. However, stiffness duration had insufficient predictive validity for initiation of definitive RA therapy and was subsequently not included in the 2010 RA classification. Recently, duration of morning stiffness has been shown to inadequately reflect the more complex experience of stiffness described by patients with RA [2–4]. This mismatch between the patient experience and the limited concept measured by morning stiffness duration represents lack of content validity and may explain why, in RA, measurement of morning stiffness duration has poor reliability and responsiveness. Halls *et al.* [5] developed, through a rigorous process of iterative qualitative data collection and analysis, a set of items for measuring patient-reported stiffness in RA. This work nicely fills an all-too-common content gap between the construct to be measured and the concept assessed by the outcome measurement instrument.

The extent to which there is correspondence between the construct intended to be measured and the concept assessed by a candidate outcome measurement instrument is termed content validity. Content validity is *sine qua non* for applicability and usefulness of outcome measurement instruments as health measurement tools [6]. Lack of content validity cannot be remediated by testing additional measurement properties [7], and recommendations from COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and OMERACT are to set aside instruments lacking adequate content validity and instead consider or develop alternate instruments [6, 8].

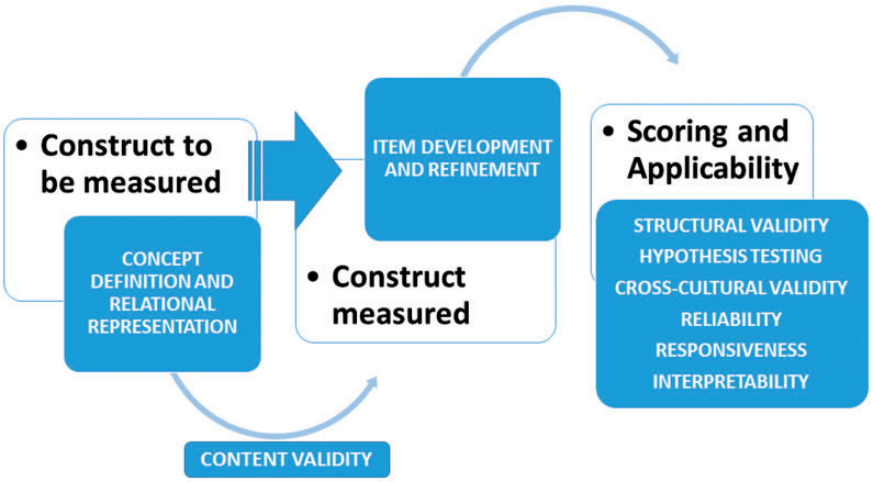
The content validity phase of outcome measurement instrument development, especially for patient-reported outcome measures, is qualitative data–driven to comprehensively define the construct in the population of interest, reveal key underpinning elements and their interrelationships [9–11], and inform draft patient-reported outcome measures that will undergo further quantitative and qualitative evaluation and refinement (Fig. 1). Inductive thematic analysis is the method of choice for exploring the full depth of a construct when no *a priori* hypotheses are available or when they are not well defined. Once hypotheses have been drafted, these can then be confirmed or challenged through additional qualitative data collection and deductive thematic analysis. The stiffness item set developed by Halls *et al.* [5] was informed by recent conceptual frameworks for stiffness in RA, which were developed based on inductive thematic analysis of interviews and focus groups with patients with RA in studies conducted by two independent groups [2, 3]. Authors then used a hypothesis-centred approach and conducted additional focus groups with patients with RA. Using deductive thematic analysis, they confirmed the relational structure of the stiffness concept. Subsequently, they used direct accounts of the patients' stiffness experience as primary data for item development, to increase item relevance for patients. An initial set of 77 stiffness items was drafted and subsequently refined with patients (through cognitive interviews) to retain highly relevant and easy-to-understand items. A final 45-item set resulted at the end of this process, and is now available for further quantitative testing and qualitative refinement so as to be applicable for different patient populations.

The next steps in developing the stiffness item set in RA would include evaluation of its measurement properties. This can be achieved through administration to patients across the continuum of stiffness and with different levels of disease activity, duration, levels of disability and demographics. Data collected would then be used to evaluate internal consistency, using classical test theory and/or item response theory approaches, to identify redundancies and gaps, and to inform its scoring algorithm. The 45-item set covers concepts of stiffness impact, timing, location, stiffness related to immobility and stiffness severity and duration [5]. These may potentially be ordered to reflect degrees of stiffness, which would make the item response theory model applicable. The measure could be further refined to retain a reduced set of the items that are most informative regarding stiffness. The final item number for a fixed short form and/or administration using a computer-adaptive testing approach would be important in determining the feasibility of the final instrument.

Further examination of reliability, responsiveness, and cross-cultural validation in RA populations would then proceed. This stiffness item set has strong evidence for content validity in RA and is highly relevant to ongoing international efforts to develop a stiffness measure applicable across rheumatic diseases. The process of validation

EDITORIAL

Fig. 1 Conceptual framework for patient-reported outcome measure development



Content validity is the essential measurement property to get right in order to ensure that the concept that needs to be measured is what is actually measured. Subsequent phases of development and testing can proceed once content validity has been established.

across rheumatic diseases would involve confirmation of content validity in the new population and additional assessments of responsiveness in other conditions. For patients, this work represents the development of a promising new stiffness measurement tool that will capture their experience of stiffness and will enable better quantification of the activity and impact of rheumatic disease. For rheumatologists, this work brings stiffness back into the spotlight and will provide clinicians with an additional tool to help gauge inflammatory disease activity and its impact on patients. For both patients and clinicians, such a measure could further facilitate shared medical decision-making.

Acknowledgements

Dr A.M.O. is supported in part by a Scientist Development Award from the Rheumatology Research Foundation.

Funding: No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: C.O.B. is funded in part through a research award from the Patient Centered Outcomes Research Institute (PCORI, ME-SC14-1402-10818) and is a member of the Executive Committee of OMERACT, which receives arm's length funding from more than 23 pharmaceutical and research companies; they receive no financial remuneration for this role. The other author has declared no conflicts of interest.

Ana-Maria Orbai[1] and Clifton O. Bingham III[1]

[1]*Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA*
Revised version accepted 26 May 2017

*Correspondence to: Ana-Maria Orbai, Johns Hopkins University School of Medicine, Division of Rheumatology, 4200 Eastern Avenue, MFL Center Tower Suite 410, Baltimore, MD, 21224, USA.
E-mail: aorbai1@jhmi.edu

References

1 Bartlett SJ, Hewlett S, Bingham CO III *et al*. Identifying core domains to assess flare in RA: an OMERACT international patient and provider combined Delphi consensus. Ann Rheum Dis 2012;71:1855–60.

2 Halls S, Dures E, Kirwan J *et al*. Stiffness is more than just duration and severity: a qualitative exploration in people with RA. Rheumatology 2015;54:615–22.

3 Orbai AM, Smith KC, Bartlett SJ, De Leon E, Bingham CO III. ''Stiffness has different meanings, I think, to everyone'': examining stiffness from the perspective of people living with RA. Arthritis Care Res 2014;66:1662–72.

4 Halls S, Sinnathurai P, Hewlett S *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 2016; Advance Access published 17 December 2017, doi: 10.3899/jrheum.161073. PubMed PMID: 27980014.

5 Halls S, Dures E, Kirwan JR *et al*. Development and testing of candidate items for inclusion in a new rheumatoid arthritis stiffness patient-reported outcome measure. Rheumatology 2017;doi: 10.1093/rheumatology/kex085. [Epub ahead of print].

6 Prinsen CA, Vohra S, Rose MR *et al*. How to select outcome measurement instruments for outcomes included in a ''Core Outcome Set'' – a practical guideline. Trials 2016;17:449.

7 U.S. Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006;4:79.

8 Boers M, Kirwan JR, Tugwell P *et al*. The OMERACT Handbook Outcome Measures in Rheumatology Clinical Trials (OMERACT). 2017; version 31 March 2017. https://www.omeract.org/pdf/OMERACT_Handbook.pdf (17 May 2017, date last accessed)

9 Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res 2009;18:1263–78.

10 Patrick DL, Burke LB, Gwaltney CJ *et al*. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. Value Health 2011;14:967–77.

11 Patrick DL, Burke LB, Gwaltney CJ *et al*. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. Value Health 2011;14:978–88.