Received: 8 April 2023 | Revised: 21 April 2023 | Accepted: 22 April 2023

DOI: 10.1002/msc.1774

WILEY

RESEARCH ARTICLE

Psychometric testing of the British-English Long-Term Conditions Job Strain Scale, Long-Term Conditions Work Spillover Scale and Work-Health-Personal Life Perceptions Scale in four rheumatic and musculoskeletal conditions

Alison Hammond[1] | Alan Tennant[2] | Angela Ching[1,3] | Jennifer Parker[1] | Yeliz Prior[1] | Monique A. M. Gignac[4,5] | Suzanne M. M. Verstappen[6,7,8] | Rachel O'Brien[9]

[1]Centre for Human Movement and Rehabilitation, School of Health and Society, University of Salford, Salford, Greater Manchester, UK

[2]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, West Yorkshire, UK

[3]King's Clinical Trials Unit, Institute of Psychiatry, King's College London, London, UK

[4]Institute of Work and Health, Toronto, Ontario, Canada

[5]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

[6]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[7]Manchester Academic Health Science Centre, Manchester, UK

[8]NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester, UK

[9]College of Health, Well Being and Life Sciences, Sheffield Hallam University, Sheffield, UK

Correspondence

Alison Hammond, Centre for Human Movement and Rehabilitation, School of Health and Society, Allerton, University of Salford, Frederick Road, Salford M6 6PU, UK.
Email: a.hammond@salford.ac.uk

Funding information

European League Against Rheumatism

Abstract

Objective: The aims were to validate linguistically British-English versions of the Long-Term Conditions Job Strain Scale (LTCJSS), Long-Term Conditions Work Spillover Scale (LTCWSS) and Work-Health-Personal Life Perceptions Scale (WHPLPS) in rheumatoid arthritis, axial spondyloarthritis, osteoarthritis and fibromyalgia (FM).

Methods: The three scales were forward translated and reviewed by an expert panel prior to cognitive debriefing interviews. Participants completed a postal questionnaire. Construct validity was assessed using Rasch analysis. Concurrent validity included testing between the three scales and work (e.g., Workplace Activity Limitations Scale [WALS]) and condition-specific health scales. Two weeks later, participants were mailed a second questionnaire to measure test-retest reliability.

Results: The questionnaire was completed by 831 employed participants: 68% women, 53.5 (SD 8.9) years of age, with condition duration 7.7 (SD 8.0) years. The LTCJSS, LTCWSS and WHPLPS Parts 1 and 2 satisfied Rasch model requirements, but Part 3 did not. A Rasch transformation scale and Reference Metric equating scales with the WALS were created. Concurrent validity was generally good (r_s = 0.41–0.85) for the three scales, except the WHPLPS Part 3. Internal consistency (Person Separation Index values) was consistent with group use in all conditions, and individual use except for the LTCWSS and WHPLSP Parts 1 and 2 in FM. Test-retest reliability was excellent, with intraclass coefficients (2,1) of 0.80–0.96 for the three scales in the four conditions.

Discussion: Reliable, valid versions of the British-English LTCJSS, LTCWSS and WHPLPS Parts 1 and 2 are now available for use in the UK.

KEYWORDS
arthritis, contextual factors, musculoskeletal, patient reported outcomes, work, work rehabilitation

1 | INTRODUCTION

Difficulties with work activities can be common amongst working people with rheumatic and musculoskeletal diseases (RMDs) (Boonen et al., 2023). Work participation (i.e., in paid employment) can be influenced by a person's health condition, a wide range of functioning and disability, personal and work-related personal and environmental contextual factors (Heerkens et al., 2017). Assessing work-related contextual factors is important to gain a greater understanding of people with RMD's ability to work (Heerkens et al., 2017; Tang et al., 2011). Amongst work-related personal factors, those associated with presenteeism (i.e., reduced productivity at work) in RMD include job strain, and work-life-health balance (Brown et al., 2023; Gignac et al., 2012). The patient reported outcome measures (PROM) to assess such factors, tested across a range of RMDs, can assist in understanding the biopsychosocial impact of RMDs on work participation as well as help plan and evaluate work interventions.

Job strain occurs when job psychological demands are high, whilst decision latitude (i.e., ability to use one's skills and make decisions) is low (Karasek et al., 1998). For those with RMD, job demands must be met within the context of their condition, which can increase stress, in turn increasing inflammation and symptom severity (Evers et al., 2014; Liu et al., 2017). If not addressed, job strain contributes to poorer health and reduced work ability (Gignac et al., 2007). The Chronic illness Job Strain Scale (CIJSS) was developed through literature review and interviews with working people with rheumatoid arthritis (RA) or osteoarthritis (OA) to identify factors contributing to perceived job stress (e.g., working with arthritis symptoms, disease uncertainty). It has good concurrent validity with the Workplace Activity Limitations Scale (WALS), a measure of presenteeism, and excellent internal consistency in working people with inflammatory arthritis (IA) (i.e., RA or psoriatic arthritis [PsA]) or OA (Gignac & Cao, 2009; Gignac et al., 2007). The Arthritis Work Spillover Scale (AWSS) assesses reciprocal demands of work on managing arthritis and arthritis interfering with work and is modelled on work-family spillover scales. It has a single factor structure and excellent internal consistency in IA (i.e., RA and PsA) and OA (Gignac et al., 2006, 2008). Working people with RMD are balancing work demands and personal lives (e.g., homemaking; caring; community, social and leisure lives) within the context of their RMD's impact and self-managing it. The Work-Health-Personal Life Perceptions Scale (WHPLPS) measures these interactions. It has good content validity, as developed through literature review and focus groups with working people with IA and OA. Factor analysis identified three distinct subscales with excellent internal consistency in IA (i.e., RA, PsA and ankylosing spondylitis [AS]) and OA (Gignac et al., 2014).

The CIJSS, AWSS and WHPLPS were developed and psychometrically tested in Canada in IA and OA, and used in studies in IA, OA, and lupus (Al Dhanhani et al., 2014, 2015; Gignac et al., 2006, 2007, 2008, 2014; Gignac & Cao, 2009). As these were developed in Canadian English, before use in the United Kingdom (UK), they should be validated linguistically (i.e., translated and culturally adapted) into British-English (a different form of the same language) and then tested psychometrically (Acquadro et al., 2004). The term 'chronic illness' is generally replaced by 'long-term conditions' in the UK. Accordingly, the CIJSS was renamed the Long-Term Conditions Job Strain Scale (LTCJSS) for the UK. The AWSS is being tested across a range of RMD, resulting in its name being changed to the Long-Term Conditions Work Spillover Scale (LTCWSS). The aims of this study were therefore to validate linguistically, investigate content validity, and evaluate the psychometrics of the British-English LTCJSS, LTCWSS and WHPLPS amongst working people with RA, axial spondyloarthritis (axSpA), lower limb OA and fibromyalgia (FM) in the UK. Testing should include both classical testing and item response theory (e.g., Rasch analysis) to establish psychometric properties (e.g., reliability and validity) (Mokkink et al., 2010).

2 | METHOD

2.1 | Design, participants, and recruitment procedures

The WORK-PROM study design used cross-cultural adaptation (Phase 1), followed by cross-sectional surveys to establish psychometric properties of the LTCJSS, LTCWSS and WHPLPS (Phase 2). The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist was followed (Gagnier et al., 2021; Mokkink et al., 2010). Phase 1 occurred in 2017 and Phase 2 from March 2018 to March 2020.

Participants were recruited from 41 secondary care and six community National Health Service Trusts' Rheumatology, Orthopaedic or Therapy out-patient clinics, with some from a University Arthritis Volunteer Register. Participants were eligible if at least 18 years old; in paid employment at least 1 day a week; currently working; and a primary diagnosis of RA; axSpA; OA (knee and/or hip); or FM. Diagnoses were confirmed by a rheumatologist for RA and axSpA; or a rheumatologist, orthopaedic surgeon, general practitioner, or extended scope physiotherapist for OA and FM. Participants needed to be able to read, write and understand British-English. Patients were ineligible if on long-term sick leave because unable to complete some of the work outcome measures requiring responses about recent ability to work. Patients were identified by

research facilitators or therapists using these criteria and given a short study explanation and information pack. The latter included a reply form, including diagnosis, employment, and sick leave status, to check eligibility criteria.

2.2 | Data collection

In Phase 1, linguistic validation, and cross-cultural adaptation were conducted to ensure that the wording in these three scales was considered comprehensible by participants (Beaton et al., 2007). Content validity (i.e., the degree to which scale content is an adequate reflection of what is being measured) was also tested during cognitive-debriefing interviews (De Vet et al., 2011). Full details of the Phase 1 method are in Supporting Information S1: File S1.

In Phase 2, for psychometric testing, participants were mailed a paper questionnaire booklet to complete at home (Test 1: T1). Two weeks after return, they were mailed a second questionnaire (Test 2: T2) to assess test-retest reliability. Following each mailing, if required, at 2 weeks, participants were sent a reminder letter, and at 4 weeks, a further reminder and questionnaire booklet.

The T1 booklet included demographic data, such as age, sex, living arrangements, education status, condition duration (of symptoms and from diagnosis), medication regimen, employment status and job title. The latter was coded into job skill-level categories (1 = elementary occupations, e.g., cleaner, refuse collector, shelf filler; 2 = requiring compulsory education/work-related training; 3 = post-compulsory education [sub-degree] or longer work experience; 4 = degree education or equivalent experience) (Office for National Statistics, 2016).

The T1 booklet also included the three scales, that is, the British-English versions of the LTCJSS, LTCWSS and WHPLPS. The LTCJSS includes 15 items, each scored 0 = not at all stressful to 4 = extremely stressful: (range 0–60), with higher scores indicating greater perceptions of job stress (Gignac et al., 2007; Supporting Information S1: File S2). The LTCWSS includes six items: three related to the impact of work on the person's health condition, and three related to the condition's impact on their work. Items are scored 0 = strongly disagree to 4 = strongly agree (range 0–24), with higher scores reflecting greater interaction between work and health. A 'not applicable' option was included in the tested version (Gignac et al., 2006; Supporting Information S1: File S3). The WHPLPS consists of 20 items scored 0 = strongly disagree to 4 = strongly agree. The three sub-scales are Part 1 Condition negatively Affects Work and personal life (CAW, eight items, range 0–32); Part 2 Work and personal life affect Condition and its management (WAC, seven items, range 0–28); and Part 3 Benefits of Work (BW, five items, range 0–20). Higher scores in Parts 1 and 2 indicate greater work-health-personal life imbalance, and in Part 3 greater benefits from work (Gignac et al., 2014; Supporting Information S1: File S4). No time period is specified in the three scales. Scoring instructions and handling of missing data for the three scales are explained in Supporting Information S1: Files S2–S4.

To test concurrent validity, a lot of work and health scales were included in the T1 questionnaire booklet. For all, a higher score indicates worse status.

Work scales: These evaluated both physical and emotional impact of conditions on work and included the British-English WALS, a measure of presenteeism, with 12 items of physical work ability (eight items); managing work demands (physically and/or mentally) (three items); and concentration at work (one item), scored 0 = no difficulty to 3 = unable to do (range 0–36) (Hammond et al., 2023). Three sub-scales of the Work Limitations Questionnaire-25 (WLQ-25) were assessed, indicating the percentage of time in the past two weeks participants had difficulty with Time Management Demands, Mental-Interpersonal Demands, and Output Demands (Lerner et al., 2001). Two forms of the Work Instability Scale (WIS) were used: the RA-WIS for RA, OA, or FM (Gilworth et al., 2003; Tang et al., 2010) and ankylosing spondylitis (AS)-WIS for axSpA (Gilworth et al., 2009). Both measure mismatch between work abilities and job demands.

Health Scales: The following were included in the T1 booklet. As some were condition-specific, four separate T1 questionnaire booklets were used, with participants completing the booklet relevant to their condition.

For RA, the RA Impact of Disease (RAID) scale, consisting of seven 0–10 numeric rating scales (NRS: e.g., pain, fatigue) scored by summing weighted scores (Gossec et al., 2011); and the Health Assessment Questionnaire (HAQ), consisting of 20 physical function items rated 0 = not at all difficult to 3 = unable to do (Kirwan & Reeback, 1986). The HAQ was scored by summing all items (0–20 = mild; 21–40 = moderate; 41–60 = severe disability) without adjustment for using aids and devices (Tennant et al., 1996; Wolfe, 2001). For axSpA, the Bath Ankylosing Spondyloarthritis Disease Activity Index (BASDAI), including items of symptom severity (e.g., pain, fatigue) (Garrett et al., 1994); and the Bath Ankylosing Spondyloarthritis Functional Index (BASFI: score range 0–10) including 10 physical function items (Calin et al., 1994). For OA, two sub-scales of the Western Ontario McMaster Universities Osteoarthritis Index (WOMAC) were included: pain (five items); and physical function (17 items), both scored 0 = none to 4 = extreme, with total scores for each sub-scale calculated (Bellamy et al., 1988). Finally, for FM, two sub-scales of the Revised Fibromyalgia Impact Questionnaire (FIQR) were included, symptoms (10 items: score range 0–50); and physical function (nine items: score range 0–30) (Bennett et al., 2009). For all four conditions, pain, fatigue, and mood 0–10 NRS were included (or extracted from health scales). Additionally, a question about perceived health status was included (Likert scale 1 = very good to 5 = very poor) for discriminant validity testing.

At Test 2, participants completed the three scales, and perceived a change in health status included for reliability testing: 'Overall, how much is your arthritis/condition troubling you now compared to when you last completed this questionnaire?' (1 = much less; 2 = less 3 = about the same; 4 = more; 5 = much more).

2.3 | Sample size

A minimum of 150 cases was needed within each condition group as Rasch analysis was used to assess construct (structural) validity (Rasch, 1980). Up to 250 samples were collected to ensure a broad spread of responses. At least 79 sets of repeated responses were needed to demonstrate that a test-retest correlation of 0.7 differs from a background correlation (constant) of 0.45, with 90% power at the 1% significance level. A test-retest reliability correlation of 0.7 is considered a minimum acceptable level (Nunnally, 1978).

2.4 | Statistical analyses

Demographic and Phase 1 item relevance scores and Phase 2 work and health scales were summarised descriptively as appropriate. RUMM 2030+ software was used for Rasch analysis (Andrich et al., 2015). As all Phase 1 items and Phase 2 scales were either ordinal or not normally distributed, non-parametric statistical tests were conducted using the Statistical Package for the Social Sciences (SPSS) v26 (IBM Corp, 2019). The following psychometric properties were assessed.

2.4.1 | Compliance

Compliance (i.e., amount of missing data) was assessed by identifying the number (%) of missing data items and also LTCJSS, LTCWSS and WHPLPS which were not scorable. Less than 3% of missing data are acceptable and more than 15% unacceptable (De Vet et al., 2011).

2.4.2 | Validity

Construct (structural) validity measures the degree to which scale scores adequately reflect the dimensionality of the construct measured, that is, do all items measure the same construct (unidimensional), and are items independent of one another. The first analytical strategy was testing the fit of the LTCJSS, LTCWSS and WHPLPS for each condition to the Rasch Model. This model specifies what should be achieved if the scale can be transformed from an ordinal- to an interval-level scale. Providing a fully integrated analytical solution, it entails tests of unidimensionality; invariance by key contextual factors (i.e., can the three scales be used to assess group differences as scale items are interpreted similarly across groups, e.g., across conditions, age groups, sex) (Teresi et al., 2000); whether or not the scale items form an appropriate probabilistic ordering consistent with the model expectations, and thus allowing the transformation of the scale to interval level measurement. Full details about the Rasch analysis are in Supporting Information S1: File S5 and elsewhere (Tennant & Conaghan, 2007).

Concurrent validity (i.e., the degree to which scale scores correlate with other relevant scales) was assessed using Spearman's correlations. We hypothesised moderate-to-strong correlations between the three scales' scores, and for each scale with the work scales, and moderate correlations for each scale with relevant condition-specific health scales. Correlations of 0.20–0.39 are considered weak, 0.4–0.59 moderate, and ≥0.6 strong (Evans, 1996).

Discriminant validity, that is, hypothesis testing that there would be significant LTCJSS, LTCWSS and WHPLPS score differences between those reporting they had very poor/poor; fair; good/very good perceived health status, was assessed using Kruskal-Wallis tests, with $p \leq 0.05$ considered significant.

2.4.3 | Reliability

Internal consistency, that is, the degree of interrelatedness between items within a scale, was assessed using Cronbach's alpha. Results ≥0.8 were deemed good to excellent: ≥0.9 is consistent with individual use; and >0.7 with group-level use (Evans, 1996). The Person Separation Index (PSI) was also calculated, for which scores >0.7 with group-level use; and ≥0.85 individual use (Tennant & Conaghan, 2007).

Test-retest reliability is the extent to which scores are the same for repeated measurements over time in those reporting that health has not changed (i.e., perceived health is 'the same' at T2). This was assessed using Spearman's correlations and intraclass correlation coefficients (ICC (2,1): two-way random consistency, average measure models). An ICC ≥0.75 is considered excellent and 0.5–0.74 moderate (Cichetti, 1994). The reliability of individual scale items was calculated using weighted kappa, with levels of agreement as 0.41–0.60 = moderate; ≥0.61 = good (Evans, 1996).

2.4.4 | Precision

Precision was assessed by calculating (a) the Standard Error of Measurement (SEM), a function of the reliability of the instrument and the standard deviation; and (b) the Smallest Detectable Difference (SDD), derived from the SEM with the formulae ($SEM \times 1.96 \times \sqrt{2}$). It is a statistical estimate of the smallest detectable difference across groups above measurement error (Donoghue, 2009; Stratford, 2004).

Floor and ceiling effects were considered present if >15% of participants achieved either the lowest or highest scores (Terwee et al., 2007). If present, these can negatively affect the quality of a scale as responsiveness (i.e., ability to detect change over time) will be limited.

3 | RESULTS

3.1 | Phase 1

Full details of the linguistic validation, cross-cultural adaptation and content validity results are in Supporting Information S1: Tables

S1–S6. In cognitive debriefing interviews ($n = 48$; participant characteristics are in Table 1), most items in all three scales were considered very or extremely relevant by participants, with no differences between conditions. Participants in job skill-level groups 1 and 2 were significantly more likely to report items as extremely relevant compared to very relevant items in groups 3 and 4. All three scales were considered comprehensive and comprehensible. Changes made by the expert panel were minor wording changes, reducing the number of words, and specifically, to the layout of the LTCJSS to include a root question, avoiding repetition of wording at the start of each item. Examples are, for the LTCJSS, 'scheduling of your job' to 'shifts or work hours' (item 3); and for the WHPLPS 'trade-offs in other areas of my life' to 'sacrifices…' (item 8). Changes reduced Flesch-Kincaid Reading Grade Level scores (i.e., reading age) in the Canadian- to the British-English versions of the: LTCJSS from 14–15 years to 11–12 years; LTCWSS from 14–15 years to 11–12 years; and the WHPLPS marginally into the 10–11 years age range.

3.2 | Phase 2

Overall, 1359 people were referred to the study. Of which, 831 returned T1 and 622 T1 and T2 booklets (Supporting Information S1: Figure S1). The response rates were secondary care 62% (696/1117), community care 53% (119/224), and volunteers 89% (16/18). Participant characteristics are shown in Table 1 and work and health scales in Table 2. Median time between tests was 36 (IQR 28–47) days.

3.2.1 | Compliance

Missing data in the scales were low. In the LTCJSS, this ranged from 0% (FM) to 0.7% (axSpA), with only 4/831 (0.48%) responses not scorable. In the LTCWSS, there were very few not scorable (only 4/831, 0.48%). However, the 'not applicable' option was selected across items in 2.19%–2.75% in RA, axSpA and OA, although only 0.03% in FM. As a result, up to 3% of LTCWSS in each of the RA, axSpA and OA groups were not scorable, although compliance was still acceptable. In the WHPLPS, missing data ranged from 0% (FM, all parts) to 0.61% (RA part 3) with only two or three from each part not scorable (Supporting Information S1: Tables S7–S9).

3.2.2 | Validity

Construct (structural) validity. The fit of the data for each scale to the Rasch model is shown in Table 3. Full details of the Rasch analysis results are in Supporting Information S1: File S5, Tables S10 and S11. The LTCJSS, LTCWSS and WHPLPS Part 1 (CAW) are all unidimensional with a good fit to the Rasch model in the four conditions, and the combined dataset (i.e., all four conditions combined, $n = 831$). The WHPLPS Part 2 (WAC) also had a good fit, except for the combined dataset, indicating it is best used in studies within conditions. However, for Part 3 (BW), whilst adequate fit in axSpA, there was only adequate fit in FM if item 16 'Work keeps me moving and active which helps my condition' was removed. There was no fit with any analysis strategy for RA, OA, or the combined dataset.

Some occasional condition-specific invariance (DIF) was observed, usually associated with sex or education. However, tests identified that differences between adjusted and unadjusted estimates were not substantive. In the combined dataset, there was slightly more DIF, particularly for the WHPLPS Part 1 (CAW). In summary, there was no invariance in the LTCJSS, LTCWSS, and WHPLPS Parts 1 (CAW) and 2 (WAC) for age, sex, condition, disease duration, educational status, employed/self-employed, or full-/part-time, supporting cross-diagnostic validity of the scales. A transformation table was created to convert raw LTCJSS, LTCWSS and WHPLPS Part 1 (CAW) scores to interval level scores, if required (Supporting Information S1: Table S12).

A Reference Metric was created allowing test equating of raw LTCJSS, LTCWSS, WHPLPS Part 1 (CAW) scores with each other and with raw WALS, RA-WIS and AS-WIS scores (Supporting Information S1: Table S13). These six scales can be considered as part of a 'Work Disturbance' domain, that is, work disruptions due to the health condition and its management. This is shown diagrammatically in Figure 1, with the WALS having the widest coverage of the Work Disturbance domain. As the RA-WIS and AS-WIS have clinically derived cut-points indicating levels of work instability, it is possible to indicate what these cut-points might also be when using the other scales. For example, a raw score of 10 on the RA-WIS (indicating moderate work instability) equates to a score of 15 on the LTCJSS, 11 on the LTCWSS and 17 on the WHPLPS Part 1 (CAW) (Supporting Information S1: Table S14).

Concurrent validity. Results are shown in Table 4 and Supporting Information S1: Table S15 for WHPLPS Part 3. The LTCJSS, LTCWSS and WHPLPS Part 1 (CAW) and Part 2 (WAC) correlated strongly with each other in the four conditions ($r_s = 0.63$–0.79), apart from the LTCWSS in OA ($r_s = 0.59$) and WHPLPS Part 2 (WAC) in FM ($r_s = 0.53$–0.58), having moderate correlations with other scales. The WHPLPS Part 3 (BW) correlated weakly with very weakly ($r_s = 0$ to -0.26) with other scales (Correlations are negative as high scores in Part 3 indicate better status).

As hypothesised, correlations with work scales were moderate to strong in the four conditions: LTCJSS $r_s = 0.61$–0.82; LTCWSS $r_s = 0.55$–0.77; WHPLPS Part 1 (CAW) $r_s = 0.57$–0.85; and WHPLPS Part 2 (WAC) $r_s = 0.41$–0.65. However, the WHPLPS Part 3 had very weak to weak correlations ($r_s = -0.01$ to -0.32). As hypothesised, correlations with health scales were generally moderate for the LTCJSS ($r_s = 0.41$–0.63), and the WHPLPS Part 1 (CAW) ($r_s = 0.43$–0.61) in all four conditions (except the latter being weak for pain, fatigue, and mood in FM, $r_s = 0.31$–0.33). For the LTCWSS, these were moderate in RA and axSpA ($r_s = 0.45$–0.57) but weak in OA and FM ($r_s = 0.22$–0.38), except for mood in OA ($r_s = 0.55$). For the WHPLPS Part 2 (WAC), correlations were moderate in RA and axSpA ($r_s = 0.45$–0.50, except for pain in RA, $r_s = 0.30$), but weak in OA

TABLE 1 Phase 1 and 2 participant characteristics.

	RA		axSpA		OA		FM	
	Phase 1 ***n* = 12**	**Phase 2** ***n* = 297**	**Phase 1** ***n* = 10**	**Phase 2** ***n* = 202**	**Phase 1** ***n* = 13**	**Phase 2** ***n* = 176**	**Phase 1** ***n* = 13**	**Phase 2** ***n* = 156**
Sex: M:F *n* (%)	5:7	77 (26.00):220 (74.00)	4:6	126 (62.40):76 (37.60)	4:9	54 (30.70):122 (69.30)	2:11	10 (6.40):146 (93.60)
Age (y.), mean (SD)	57.33 (6.77)	53.52 (8.93)	33.00 (14.62)	46.80 (10.25)	55.92 (6.70)	55.99 (7.52)	39.69 (9.11)	45.71 (10.05)
Job skill level, *n* (%)								
- 1 and 2	3	151 (51.00)	4	66 (32.70)	8	86 (48.90)	7	95 (61.00)
- 3 and 4	9	143 (48.00)	5	136 (67.30)	5	89 (50.60)	6	61 (39.00)
- Missing	-	3 (1.00)	1	-	-	1 (0.60)	-	-
Disease duration (y.), mean (SD)	18.08 (11.93)	7.68 (8.02)	12.70 (9.78)	12.25 (10.35)	12.35 (10.60)	4.87 (6.80)	5.38 (3.55)	2.99 (4.17)
- Duration ≤ 2y, *n* (%)	-	89 (30.00)	-	37 (18.40)	-	93 (52.80)	-	99 (63.50)
Phase 2 only:								
Symptom duration (y.), mean(SD)	9.34 (8.56)		18.89 (11.69)		7.87 (8.43)		8.36 (7.16)	
- Duration ≤ 2y, *n* (%)	50 (16.90)		3 (1.50)		38 (21.68)		22 (14.10)	
Living with spouse/ family/significant other, *n* (%)	243 (81.80)		181 (89.60)		146 (83.00)		139 (89.10)	
Children ≤18y. living at home, *n* (%)	69 (23.20)		69 (34.20)		32 (18.40)		56 (35.90)	
Educational level (ISCED), *n* (%)								
- No formal education qualifications	28 (9.40)		14 (6.90)		17 (9.70)		7 (4.50)	
- Secondary/non tertiary	150 (50.50)		100 (49.50)		93 (52.90)		76 (48.70)	
- Tertiary	117 (39.40)		87 (43.10)		62 (35.30)		73 (46.80)	
- Missing	2 (0.70)		1 (0.50)		4 (2.30)		-	
Hours worked, median (IQR)	35.00 (21.25–37.50)		37.50 (32.25–39.00)		35.50 (24.00–37.50)		29.56 (10.17)	
Self-employed, *n* (%)	64 (21.50)		35 (17.30)		22 (12.50)		18 (11.50)	

TABLE 1 (Continued)

	RA		axSpA		OA		FM	
	Phase 1 *n* = 12	**Phase 2** *n* = 297	**Phase 1** *n* = 10	**Phase 2** *n* = 202	**Phase 1** *n* = 13	**Phase 2** *n* = 176	**Phase 1** *n* = 13	**Phase 2** *n* = 156
Medication regimen, *n* (%)								
- None	2 (0.70)		19 (9.40)		34 (19.30)		0	
- NSAIDS +/− analgesics	11 (3.70)		4 (2.00)		121 (68.80)		14 (9.00)	
- Steroids +/- NSAIDS	6 (2.00)		52 (25.70)		10 (5.70)		6 (3.80)	
- Single DMARD	104 (35.00)		10 (5.00)		-		-	
- Combination DMARD	98 (33.00)		2 (1.00)		-		-	
- Biologic/biosimilar	67 (22.6)		114 (56.40)		-		-	
- FM: Neuropathic analgaesics (e.g., gabapentin/ pregabalin/ amitriptyline)	-		-		-		99 (63.50)	
- FM: Opiate medication	-		-		-		12 (7.70)	
Time between test 1 and test 2: days, median (IQR)	40 (34.00–48.00) (*n* = 221)		38 (29.00–49.25) (*n* = 154)		30 (23.75–37.00) (*n* = 130)		33 (26.50–45.00) (*n* = 117)	

Abbreviations: axSpA, axial spondyloarthritis; DMARD, disease modifying anti-rheumatic drug; FM, fibromyalgia; IQR, inter-quartile range; NSAID, non-steroidal anti-inflammatory drug; OA, osteoarthritis; RA, rheumatoid arthritis; SD, standard deviation.

TABLE 2 Participants' work and health scales.

Median (IQR)	RA (*n* = 297)	axSpA (*n* = 202)	OA (*n* = 176)	FM (*n* = 156)
Work scales				
LTCJSS (0–60)	22.00 (12.00–36.00)	15.00 (8.00–29.00)	24.00 (12.00–36.00)	43.00 (31.00–51.00)
LTCWSS (0–24)	13.00 (8.00–16.75)	11.00 (6.00–15.00)	13.00 (9.00–16.00)	16.00 (13.00–19.00)
WHPLPS				
- 1. CAW: 0–32	20.00 (14.00–25.00)	16.00 (10.00–21.25)	20.00 (14.00–24.00)	26.00 (21.00–29.00)
- 2. WAC: 0–28	14.00 (9.00–19.00)	13.00 (7.75–17.00)	16.00 (10.00–19.00)	21.00 (17.00–24.75)
- 3. BW:0–20	15.00 (13.00–17.00)	14.00 (11.00–17.00)	15.00 (12.00–17.00)	14.00 (11.00–17.00)
WALS (0–36)	9.00 (5.00–14.00)	6.00 (3.00–11.00)	10.00 (6.00–14.00)	16.00 (12.00–19.00)
WLQ-25 (0–100)				
- Time management demands	30.00 (10.00–55.00)	25.00 (5.00–50.00)	25.00 (10.00–50.00)	60.00 (40.00–80.00)
- Mental interpersonal demands	16.67 (5.55–36.11)	13.88 (2.78–29.95)	16.66 (5.56–34.03)	44.44 (27.78–61.11)
- Output demands	20.00 (5.00–41.00)	10.00 (0–30.00)	20.00 (5.0–43.75)	45.00 (25.00–65.00)
Work instability scale (RA-WIS: 0–23 RA, OA, FM; AS-WIS 0–20)	13.00 (7.50–18.00)	10.50 (4.00–15.00)	13.00 (8.00–17.00)	18.00 (15.00–20.00)
Health scales				
Perceived severity health last month (1–5)	3.00 (2.00–3.00)	2.00 (2.00–3.00)	3.00 (3.00–3.00)	4.00 (3.00–4.00)
Pain NRS (0–10)	5.00 (3.00–7.00)	4.80 (1.60–6.90)	-	7.00 (6.00–8.00)
Fatigue NRS (0–10)	6.00 (4.00–8.00)	6.00 (2.40–7.50)	6.00 (4.00–8.00)	8.00 (7.00–9.00)
Mood NRS (0–10)	4.00 (2.00–7.00)	4.00 (2.00–6.00)	5.00 (3.00–7.00)	6.50 (5.00–8.00)
RA				
- RAID (0–10)	4.85 (3.15–6.47)	-	-	-
- HAQ20 (0–60)	9.00 (3.00–18.00)	-	-	-
axSpA				
- BASDAI (0–10)	-	3.91 (1.95–5.85)	-	-
- BASFI (0–10)	-	2.96 (1.31–5.33)	-	-
OA				
WOMAC				
- Physical function (0–68)	-	-	30.50 (21.00–41.00)	-
- Pain (0–20)	-	-	10.00 (7.00–13.00)	-
FM				
FIQR (normalised scores)				
- Symptoms (0–50)	-	-	-	34.50 (28.13–39.00)
- Function (0–30)	-	-	-	19.33 (14.67–22.67)

Abbreviations: axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disability Index; BASFI, Bath Ankylosing Spondylitis Function Index; FIQR, Fibromyalgia Impact Questionnaire – Revised; FM, fibromyalgia; HAQ, Health Assessment Questionnaire; LTCJSS, Long-Term Conditions Job Strain Scale; LTCWSS, Long-Term Conditions Work Spillover Scale; NRS, numeric rating scale; OA, osteoarthritis; RA, rheumatoid arthritis; RAID, RA Impact Disease; WALS, Workplace Activity Limitations Scale; WHPLPS, Work-Health-Personal Life Perceptions Scale (1. CAW, Condition negatively Affects Work and personal life; 2. WAC, Work and personal life affect Condition and its management; 3. BW, Benefits of Work); WLQ-25, Work Limitations Questionnaire–25; WOMAC, Western Ontario McMaster Universities Osteoarthritis Index.

(r_s = 0.23–0.28), except for mood (r_s = 0.46); and weak to moderate in FM (r_s = 0.38–0.49). The WHPLPS Part 3 (BW) correlated very weakly (r_s = 0.04 to −0.32). Correlations were significant at $p < 0.01$ (except for WHPLPS Part 3).

Discriminant validity. There were significant differences between the three levels of perceived disease severity across all four conditions for the LTCJSS, LTCWSS, and WHPLPS Parts 1 (CAW) and 2 (WAC) but not Part 3 (BW) (Supporting Information S1: Table S16).

TABLE 3 Fit of the LTCJSS, LTCWSS and WHPLPS (parts 1–3) to the Rasch model: construct (structural) validity.

Scale/diagnosis	Residuals (SD)		Chi-square		Reliability		Dimensionality			Latent correlation[a]
	Item	Person	Value (df)	*p*	PSI	α	% *t*-tests (LCI)	DIF	ECV	
LTCJSS										
- RA	0.40	0.93	55.20 (49.00)	0.25	0.92	0.94	3.60	Sex, education	0.97	0.94
- axSpA	0.38	0.86	41.80 (43.00)	0.52	0.92	0.96	3.48	None	0.98	0.97
- OA	0.37	0.86	48.50 (48.00)	0.45	0.92	0.93	5.30 (2.00)	None	0.97	0.95
- FM	0.32	0.83	55.10 (41.00)	0.07	0.90	0.93	2.56	None	0.96	0.93
Across all 4 conditions	0.78	0.93	72.50 (52.00)	0.03	0.93	0.95	4.20	Sex, education	0.97	0.96
LTCWSS										
- RA	0.15	0.97	27.10 (16.00)	0.04	0.88	0.89	3.10	None	0.99	0.99
- axSpA	0.02	0.90	14.10 (14.00)	0.45	0.85	0.86	1.02	None	0.98	0.94
- OA	0.05	0.80	15.70 (14.00)	0.33	0.86	0.85	3.50	None	0.99	0.97
- FM	0.27	0.79	7.50 (12.00)	0.82	0.82	0.80	2.59	None	0.96	0.90
Across all 4 conditions	0.17	0.86	20.10 (16.00)	0.21	0.88	0.88	2.10	None	0.99	0.99
WHPLPS: Part 1 (CAW)										
- RA	1.48	1.27	79.90 (72.00)	0.24	0.87	0.89	6.40 (3.90)	None	-	-
- axSpA	1.30	1.16	17.40 (18.00)	0.50	0.88	0.90	5.00	None	0.97	-
- OA	0.97	1.30	36.30 (32.00)	0.28	0.87	0.88	7.40 (4.20)	None	-	-
- FM	0.92	1.09	12.10 (12.00)	0.44	0.73	0.80	1.90	None	0.97	-
Across all 4 conditions	1.21	0.88	35.30 (24.00)	0.06	0.83	0.87	2.30	None	0.95	0. 89
WHPLPS part 2 (WAC)										
- RA	1.13	1.02	35.50 (20.00)	0.02	0.85	0.87	3.40	None	0.95	0.93
- axSpA	1.38	1.32	36.40 (21.00)	0.02	0.87	0.88	6.40 (3.40)	None	-	-
- OA	1.93	1.28	22.00 (15.00)	0.11	0.90	0.89	7.40 (4.20)	None	0.96	-
- FM	2.11	0.96	10.40 (8.00)	0.24	0.72	0.77	1.30	None	0.88	-
Across all 4 conditions	6.84	0.93	29.80 (20.00)	0.07	0.83	0.64	2.20	Condition, sex, hours	0.71	0.77
WHPLPS Part 3 (BW)										
- RA	3.49	1.09	21.60 (9.00)	0.01	0.72	0.66	0	None	0.77	0.76
- axSpA	2.18	1.52	61.40 (45.00)	0.05	0.82	0.87	5.50 (2.40)	Sex	-	-
- OA	1.74	0.89	32.00 (11.00)	0.001	0.75	0.75	2.90	Sex, education	0.87	0.80
- FM	0.92[a]	1.26	21.60 (28.00)	0.80	0.82	0.89	3.90[b]	None	-	-
Across all 4 conditions	2.99	0.76	81.70 (12.00)	0.00	0.74	0.76	3.30	None	0.93	0.88
Ideal values	**<1.4**	**<1.4**		**>0.01**	**>0.7**	**>0.7**	**<5%**		**>0.9**	**>0.9**

Abbreviations: α, Cronbach's alpha; axSpA, axial spondyloarthritis; DIF, Differential Item Functioning; ECV, Explained Common Variance; FM, fibromyalgia; LCI, Lower Confidence Interval; LTCJSS, Long-Term Conditions Job Strain Scale; LTCWSS, Long-Term Conditions Work Spillover Scale; OA, osteoarthritis; PSI, Person Separation Index; RA, rheumatoid arthritis; SD, Standard Deviation; WHPLPS, Work-Health-Personal Life Perceptions Scale (1. CAW, Condition negatively Affects Work and personal life; 2. WAC, Work and personal life affect Condition and its management; 3. BW, Benefits of Work).

[a]Involves item deletion.

[b]Low Power.

3.2.3 | Reliability

Internal consistency. Cronbach's alpha values were mostly good to excellent (0.76–0.96), consistent with group-level use, except for the WHPLPS Part 2 (WAC) in the combined dataset, and Part 3 (BW) in RA (Table 3). The PSI values were also good (0.72–0.92).

Test-retest reliability. At T2, 356/622 (57%) reported their condition was 'the same' as at T1 and included in analyses. For all four

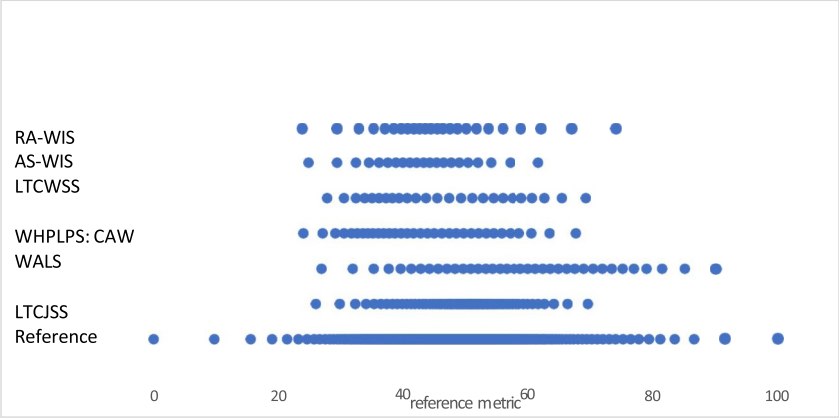

FIGURE 1 Calibration of the 'work disturbance' scales: equated tests on the reference metric. AS-WIS, Ankylosing Spondylitis Work Instability Scale; LTCJSS, Long-Term Condition Job Strain Scale; LTCWSS, Long-Term Conditions Work Spillover Scale; RA-WIS, Rheumatoid Arthritis Work Instability Scale; WALS, Workplace Activity Limitations Scale; WHPLPS: CAW, Work-Health-Personal Life Perceptions Scale (Part 1 Condition negatively affects Work and personal life).

conditions, correlations between T1 and T2 scores were strong to very strong for the LTCJSS, LTCWSS, and WHPLPS (Parts 1,2 and 3) (r_s = 0.63–0.92). ICC (2,1) were excellent at 0.80–0.96 (Table 5). Item reliability was moderate to good (Supporting Information S1: Tables S17–S19).

3.2.4 | Precision

Precision. The SEM and SDD scores for the scales are shown in Table 5.

Floor and ceiling effects. Between 0% and 9.90% of participants scored either the lowest or highest scores on the three scales, that is, within acceptable limits (<15%) (Supporting Information S1: Table S20).

4 | DISCUSSION

Linguistically validated British-English versions of the LTCJSS, LTCWSS and WHPLPS are now freely available for use in the UK (Supporting Information S1). This study provides new evidence that the LTCJSS, LTCWSS and WHPLPS Parts 1 (CAW) and 2 (WAC) have good psychometric properties in RA, axSpA, OA and FM in the UK.

The scales demonstrated good content validity from the patient perspective, as items were considered very or extremely relevant by participants across conditions. Participants in job skill-level groups 1 and 2 rated more items as 'extremely relevant', compared to 'very relevant' in groups 3 and 4, indicating that job strain, spillover and work-health-life balance are issues impacting even more on less affluent working people's jobs and lives. This indicates the importance of exploring these issues with working people with RMDs. Linguistic and cross-cultural validity were aided by ensuring input from patient research partners as well as the scales' developer approving changes. Wording changes helped reduce the reading age of scales by two to 3 years to 10–12 years. Most adults in England have a reading age of 11–14 years, meaning these scales should be understandable for most people (Health Education England). However, 15% of adults in the UK have literacy levels at or below 11 years of age, meaning alternate scales would be needed for those with poor literacy skills (National Literacy Trust, 2017). Participants' comments indicated the scales were thought-provoking, helping in re-appraising the impact of their RMD on their work and lives, and some indicated they were prompted to consider actions to reduce work stress and achieve better work-life balance (WLB). '*It would be good to have to do every year. If you have a long-term condition, you just get on with it! It helps to re-appraise or revisit your situation and think, is there anything more I need to do, as things change…. You can be so busy getting on with life you aren't thinking about the impact of your condition*' (Supporting Information S1). This suggests that completing the scales may help with addressing work and health problems and contribute to patient activation.

This is the first study to examine the construct (structural) validity of these British-English scales in RA, axSpA, OA and FM, demonstrating the fit to the Rasch model, except for the WHPLPS Part 3 (BW), which only demonstrated fit in axSpA. These scales (except for the WHPLPS Part 3) were unidimensional, meaning that raw scores can be summed or (Rasch) standardised scores used. A Rasch transformation table is available to covert raw to interval scores, as is a Reference Metric allowing test equating between the LTCJSS, LTCWSS and WHPLPS Part 1 and the WALS, RA- and AS-WIS, that is, six scales all having an underlying trait of 'Work Disturbance'. The LTCJSS, LTCWSS and WHPLPS Parts 1 and 2 generally demonstrated good concurrent validity with work and health scales, although weaker with health scales in the LTCWSS and WHPLPS Part 2 (WAC) in OA and FM. Potentially, work may have

TABLE 4 Concurrent validity of the LTCJSS, LTCWSS, WHPLPS with work and health measures (RA = 297; axSpA = 202; OA = 176; FM n = 156).

	LTCJSS (r_s)				LTCWSS (r_s)				WHPLPS: Part 1 (CAW) (r_s)				WHPLPS: Part 2 (WAC) (r_s)			
	RA	AxSpA	OA	FM	RA	AxSpA	OA	FM	RA	AxSpA	OA	FM	RA	AxSpA	OA	FM
LTCWSS	0.74	0.77	0.72	0.68	-	-	-	-	-	-	-	-	-	-	-	-
WHPLPS part 1 (CAW)	0.79	0.78	0.82	0.78	0.74	0.79	0.66	0.64	-	-	-	-	-	-	-	-
WHPLPS part 2 (WAC)	0.66	0.68	0.63	0.58	0.68	0.69	0.59	0.53	0.67	0.68	0.69	0.57	-	-	-	-
Work scales																
WALS	0.71	0.76	0.72	0.64	0.66	0.77	0.58	0.55	0.71	0.79	0.71	0.65	0.45	0.62	0.47	0.42
WLQ-25																
- Time management demands	0.63	0.75	0.64	0.74	0.63	0.69	0.59	0.60	0.63	0.73	0.57	0.59	0.47	0.63	0.45	0.53
- Mental–interpersonal demands	0.70	0.69	0.72	0.68	0.64	0.66	0.62	0.64	0.68	0.76	0.68	0.66	0.56	0.55	0.54	0.54
- Output demands	0.69	0.70	0.61	0.65	0.70	0.69	0.57	0.64	0.72	0.75	0.59	0.65	0.57	0.59	0.47	0.41
WIS	0.77	0.82	0.78	0.69	0.74	0.77	0.68	0.63	0.78	0.85	0.76	0.66	0.60	0.65	0.54	0.49
Health scales																
Pain NRS	0.47	0.47	-	0.41	0.46	0.45	-	0.22	0.45	-	-	0.33	0.30	-	-	0.24
Fatigue NRS	0.62	0.58	0.61	0.46	0.57	0.51	0.38	0.33	0.58	0.59	0.57	0.31	0.50	0.49	0.39	0.42
Mood NRS	0.56	0.55	0.62	0.44	0.49	0.49	0.55	0.26	0.52	0.59	0.49	0.31	0.50	0.38	0.46	0.45
RA																
- RAID	0.63	-	-	-	0.56	-	-	-	0.60	-	-	-	0.46	-	-	-
- HAQ20	0.54	-	-	-	0.46	-	-	-	0.59	-	-	-	0.45	-	-	-
axSpA																
- BASDAI	-	0.58	-	-	-	0.52	-	-	-	0.58	-	-	-	0.49	-	-
- BASFI	-	0.54	-	-	-	0.54	-	-	-	0.61	-	-	-	0.48	-	-
OA																
- WOMAC pain	-	-	0.49	-	-	-	0.23	-	-	-	0.44	-	-	-	0.28	-
- WOMAC physical function	-	-	0.46		-	-	0.24	-	-	-	0.43	-	-	-	0.23	-
FM																
- FIQR symptoms	-	-	-	0.62	-	-	-	0.37	-	-	-	0.49	-	-	-	0.49
- FIQR function	-	-	-	0.53	-	-	-	0.33	-	-	-	0.45	-	-	-	0.38

Abbreviations: axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disability Index; BASFI, Bath Ankylosing Spondylitis Function Index; FIQR, Fibromyalgia Impact Questionnaire – Revised; FM, fibromyalgia; HAQ, Health Assessment Questionnaire; LTCJSS, Long-Term Conditions Job Strain Scale; LTCWSS, Long-Term Conditions Work Spillover Scale; NRS, numeric rating scale; OA, osteoarthritis; RA, rheumatoid arthritis; RAID, Rheumatoid Arthritis Impact of Disease; r_s, Spearman's correlations—all correlations are significant at the $p \leq 0.01$; WALS, Workplace Activity Limitations Scale; WHPLPS, Work-Health-Personal Life Perceptions Scale (CAW, Disease affects Work; WAC, Work and personal life affect disease); WIS, Work Instability Scale; WLQ-25, Work Limitations Questionnaire-25; WOMAC, Western Ontario McMaster Universities Osteoarthritis Index.

less impact on OA and FM (as opposed to OA and FM affecting work, which was high), compared to RA and axSpA, than hypothesised, and warrants further investigation. It was also notable in OA that, across the three scales, fatigue and mood correlated more than pain and function. Internal consistency for the three scales was also good and comparable to findings in RA and OA in Canada (Gignac et al., 2006, 2007, 2014). As the PSI values for the three scales were above 0.7, all can be used for group measurement in RA, axSpA, OA and FM. The PSI values also indicated that the scales are suitable for individual use, as values were above 0.85, except for the LTCWSS and WHPLSP Parts 1 and 2 in FM. In FM, within-person changes for these two scales should be interpreted with caution as PSI values were sufficient for group use only. The study also provided the first evidence for test-retest reliability for the LTCJSS, LTCWSS and WHPLPS in each of the four RMDs. In the LTCWSS, whilst compliance was acceptable, removing the 'not applicable' option would force a response (The most likely being 'strongly disagree' as, if an item is not applicable, it is not problematic). The 'not applicable' column was

TABLE 5 Test-retest reliability and precision of the LTCJSS, LTCWSS and WHPLPS.

	n for test-retest[a]	T1 score median (IQR)	T2 score median (IQR)	Correlation T1 T2 (r_s)	ICC (2,1) (95% CI)	SEM	SDD
LTCJSS (0–60)							
- RA	136	18.00 (10.25–31.00)	16.00 (10.25–27.00)	0.87**	0.93 (0.90, 0.95)	2.97	8.38
- axSpA	99	13.00 (7.25–24.00)	12.00 (5.00–23.00)	0.92**	0.96 (0.94, 0.97)	1.83	5.18
- OA	79	24.00 (10.00–35.00)	17.00 (9.00–33.00)	0.86**	0.92 (0.88, 0.95)	3.35	9.46
- FM	54	41.00 (26.75–51.25)	40.50 (26.75–51.00)	0.86**	0.93 (0.88, 0.96)	2.83	7.98
LTCWSS (0–24)							
- RA	130	13.00 (8.00–16.00)	12.00 (8.00–15.00)	0.82**	0.91 (0.87, 0.93)	1.77	4.99
- axSpA	97	10.00 (6.00–13.25)	9.00 (6.00–14.00)	0.80**	0.88 (0.82, 0.92)	1.76	4.97
- OA	76	13.00 (8.75–16.00)	12.00 (7.25–16.00)	0.81**	0.91 (0.86, 0.94)	1.70	4.80
- FM	54	16.00 (13.00–18.00)	17.00 (13.00–19.00)	0.63**	0.78 (0.62, 0.87)	1.91	5.38
WHPLPS Part 1 (CAW: 0–32)							
- RA	136	19.00 (13.25–23.00)	18.00 (13.00–23.00)	0.81**	0.91 (0.87, 0.94)	2.03	5.72
- axSpA	100	14.00 (8.25–21.00)	14.00 (8.00–20.75)	0.89**	0.90 (0.85, 0.93)	2.11	5.95
- OA	79	18.00 (13.00–24.00)	19.00 (11.00–23.00)	0.84**	0.93 (0.88, 0.95)	1.69	4.76
- FM	54	26.00 (21.00–29.00)	25.50 (19.00–29.00)	0.81**	0.92 (0.86, 0.95)	1.21	3.40
WHPLPS Part 2 (WAC: 0–28)							
- RA	136	14.00 (9.00–19.00)	14.00 (9.00–18.00)	0.82**	0.91 (0.87, 0.93)	1.72	4.86
- axSpA	99	12.00 (7.00–17.00)	13.00 (6.00–18.00)	0.77**	0.80 (0.71, 0.87)	2.68	7.56
- OA	79	15.00 (7.00–20.00)	14.00 (8.00–19.00)	0.86**	0.93 (0.90, 0.96)	1.23	3.46
- FM	54	21.00 (16.00–15.00)	22.00 (16.75–26.00)	0.73**	0.88 (0.79, 0.93)	1.69	4.77
WHPLPS Part 3 (BW: 0–20)							
- RA	136	15.00 (13.00–16.00)	15.00 (13.00–18.00)	0.64**	0.80 (0.56, 0.91)	1.81	5.12
- axSpA	99	14.00 (11.00–18.00)	14.00 (12.00–19.00)	0.71**	0.83 (0.75, 0.89)	1.41	3.97
- OA	79	14.00 (12.00–15.75)	14.50 (12.00–17.00)	0.72**	0.84 (0.75, 0.90)	1.45	4.10
- FM	54	14.00 (11.75–17.00)	14.00 (11.75–17.25)	0.82**	0.87 (0.78, 0.92)	1.25	3.54

Abbreviations: axSpA, axial spondyloarthritis; FM, fibromyalgia; ICC, intra-class correlation coefficient; IQR, inter-quartile range; LTCJSS, Long-Term Conditions Job Strain Scale; LTCWSS, Long-Term Conditions Work Spillover Scale; OA, osteoarthritis; RA, rheumatoid arthritis; SDD, Smallest Detectable Difference; SEM, Standard Error of Measurement; WHPLPS, Work-Health-Personal Life Perceptions Scale (1. CAW, Condition negatively Affects Work and personal life; 2. WAC, Work and personal life affect Condition and its management; 3. BW, Benefits of Work).

[a]Participants indicating perceived health 'about the same' at T1 and T2, who had scores available at both time points.

**Correlation significant at $p \leq 0.01$.

therefore removed. As a 'not applicable' option was unavailable in the LTCJSS and WHPLPS, which both had good compliance, its removal in the LTCWSS is unlikely to be problematic.

However, the WHPLPS Part 3 (BW) did not fit the Rasch model (except in axSpA) and had poor concurrent and discriminant validity in these four RMDs. As a result, it is recommended that this sub-scale is omitted for research studies, as it does not assess a consistent or single construct. Phase 1 participants reported particularly liking the WHPLPS Part 3 (BW) with its positive focus on benefits of working, in contrast to the negative focus in Parts 1 and 2 and the other scales. Part 3 could still be beneficial in clinical use to prompt reflection and discussion about individual items but not scored.

The LTCWSS and WHPLPS Parts 1 and 2 partly overlap conceptually, as both measure the impact of the condition on work, and vice versa. Of the two, the WHPLPS is preferable for clinical and research use as it includes personal life. However, if a shorter scale is required, the LTCWSS is available.

There are other job strain and WLB scales available. In industry, the most widely used job strain scale is the Job Content Questionnaire (JCQ), which measures social and psychological characteristics of jobs (decision latitude, psychological job demands, social support [from employer and co-workers], physical demands, job insecurity, emotional demands, and organisational level) (Job Content Questionnaire Centre, 2021; Karasek et al., 1998). Although there is some

overlap in item content, the LTCJSS focuses on job strain from the individual's perspective in the context of their health condition rather than assessing the job, making the LTCJSS more applicable to clinical studies. One frequently used WLB scale in research was developed by Hayman (2005), based on an earlier scale (Fisher-McAulley et al., 2003). This has three-parts: work interferes with personal life, personal life interferes with work, and work/personal life enhancement. This, like the WHPLPS, takes a broader approach than many WLB scales (which focus only on work and family life) as it includes non-work activities, making it relevant for a wider range of workers. However, it does not consider the impact of health on work and personal life, that is, the focus of the WHPLPS.

4.1 | Strengths, limitations and future research

Relatively large samples of working people with RA, axSpA, OA and FM were recruited across the UK, meaning that results are representative of those accessing secondary and community care. This was not a community-based study, so those with OA and FM, especially, may have worse health status than if a community sample was recruited. In FM, very few men were recruited, and the test-retest sample was smaller than required. In general, participants had either longer symptoms or disease durations, and may represent those managing to stay in employment. For the purposes of psychometric testing, an appropriate range of participants was recruited.

Responsiveness (i.e., longitudinal validity) still needs to assess and minimal clinically important differences (MCID) established. Further testing in other RMDs is required. The scales may be suitable for other long-term conditions following testing as items are not condition-specific, apart from item 1 in the LTCJSS, with symptom examples of pain and fatigue. These are common in other health conditions, but examples could be changed as applicable.

4.2 | Conclusion

Overall, there is good validity and reliability of the British-English LTCJSS, LTCWSS and WHPLPS Parts 1 and 2 in working people with RA, axSpA, OA, or FM in the UK, but not for the WHPLPS Part 3 as a measure assessing a single construct. The latter is therefore not suitable for use in research. The three scales meet most recommendations of the COSMIN checklist for methodological quality and reporting (Gagnier et al., 2021; Mokkink et al., 2010). Accordingly, the scales (excluding WHPLPS Part 3) can be used in the UK in these four RMDs in clinical practice and research. Transformations to metric scales are available for calculations of change and other parametric procedures, distribution permitting.

AUTHOR CONTRIBUTIONS

Alison Hammond, Alan Tennant and Yeliz Prior contributed to the study conception and design. Phase 1: Alison Hammond and Yeliz Prior conducted data collection and analysis. Alison Hammond, Alan Tennant, Monique A. M. Gignac, Yeliz Prior, Suzanne M. M. Verstappen and Rachel O'Brien were members of the Expert Panel. Phase 2: Material preparation and data collection were performed by Alison Hammond, Angela Ching and Jennifer Parker. Analyses were performed by Alan Tennant (Rasch analysis) and Alison Hammond (classical testing). The first draft of the manuscript was prepared by Alison Hammond. All authors contributed to previous versions of the manuscript. All authors have read and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors wish to thank: all the study participants for their time in completing the questionnaires. In Phase 1, the expert panel members for their time in supporting the linguistic validation process: John Grogan (JG: translator); Anita Prince (AP) and Stephen Kay (SK): patient research partners; Tracy White (TW: Senior Occupational Therapist, Wrightington Hospital); and Yvonne Hough (YH: Senior Occupational Therapist, St Helens Hospital), clinical advisors. In Phase 2: all the Local Collaborators (LC), rheumatology consultants, research facilitators, rheumatology nurses, occupational therapists and physiotherapists assisting with participant identification at the contributing NHS Trusts:

ENGLAND: Yvonne Hough (LC), St Helens and Knowsley Teaching Hospitals NHS Trust; Tracy White (LC), Wrightington, Wigan and Leigh NHS Foundation Trust; Karen Crosby (LC), JoAnn Nicholson, Susannah Glasgow, Manchester University NHS Foundation Trust; Nicky Walker (LC), Mid Cheshire Hospitals NHS Foundation Trust; Sarah Wilson (LC), Sheffield Teaching Hospitals NHS Foundation Trust; Carol Graham (LC), Midlands Partnership NHS Foundation Trust; Anne Boulton (LC), Northumbria Healthcare NHS Foundation Trust; Clare Webb (LC), University Hospitals of Derby and Burton NHS Foundation Trust; Anne Bontoft (LC), Northern Lincolnshire and Goole NHS Foundation Trust; Deborah Wilson (LC), Phil Buckley, Sherwood Forest Hospitals NHS Foundation; Dr Roshan Amarasena (LC), Jayne Edwards, Theresa Grant, Lisa Burgess, The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust; Sandi Derham (LC), Royal United Hospitals Bath NHS Foundation Trust; Sarah Small (LC), James Paget University Hospitals NHS Foundation Trust; Christina MacLeod (LC), Kevin Spear, Emily Porter, Hampshire Hospitals NHS Foundation Trust; Dr Shweta Bhagat (LC), Julie Blundell, Nyssa Muskett, West Suffolk NHS Foundation Trust; Sathish Govindarajalu (LC), Matthew Pearson, Samantha Nunn, Cambridgeshire Community Services NHS Trust; Dr David Coady (LC), Rona Ymballa, City Hospitals Sunderland NHS Foundation Trust; Susan Ellis (LC), East Lancashire Hospitals NHS Trust; Peter Swan (LC), Stephanie Howard, Wendy Neale, Norfolk Community Health and Care NHS Trust; Helen Jeffrey (LC), Jude Prince, Countess of Chester Hospital NHS Foundation Trust; Mr Kishore Mamidi (LC), Dr Marwan Bukhari (LC), Kathryn Allison, Jackie Toomey, Lynda Fothergill, University Hospitals of Morecambe Bay NHS Foundation Trust; Dr Karl Gaffney (LC), James Kennedy, Celia Woodhouse, Norfolk and Norwich University Hospitals NHS Foundation Trust; Dr Pamela Peterson (LC), Susan Pugmire, Gateshead Health NHS Foundation Trust; Dr Hanadi Sari-Kouzel (LC), Marina Oprea, Greta Van

Duyvenvoorde, Allison Clarke, Blackpool Teaching Hospitals NHS Foundation Trust; Dr Suzanne Lane (LC), Cathleen Chabo, Sue Brixey, Ipswich Hospital NHS Trust; Dr Imran Riaz (LC), Ellie Gilham, Jayne Brown, The Queen Elizabeth Hospital King's Lynn NHS Foundation Trust; Suzannah Pegler (LC), Great Western Hospitals NHS Foundation Trust; Louise Hollister (LC), Dawn Simmons, Weston Area Health NHS Trust; Sue Smolen (LC), Mid Essex Hospital Services NHS Trust; David Sweeting (LC), East Coast Community Healthcare; Emma McLoughlin (LC), Anna Thornhill, Charlotte Dando, Solent NHS Trust; Dr Sophia Naz (LC), Lorraine Lock, Northern Care Alliance; Jonathan Price (LC), Joanne Holt, Birmingham Community Healthcare NHS Foundation Trust; Alison Bradshaw (LC), The Walton Centre NHS Foundation Trust; Dr Dobrina Hull and Dr Prahbu Gandhimani (LC), Kingston Hospital NHS Foundation Trust; Fiona Wright (LC), North Bristol NHS Trust; Dr Michael Green (LC), Samantha Roche, Holly Hancock, York Teaching Hospitals NHS Foundation Trust; Dr Cathy Lawson (LC), Pauline Fitzgerald, Lynsey Hall, Harrogate & District NHS Foundation Trust; Liz Lowe (LC), Jacqueline McCormick, Tameside and Glossop Integrated Care NHS Foundation Trust; Lee Hawthorn (LC), Bridgewater Community Healthcare NHS Foundation Trust; Dr Jill Firth (LC), Katherine Kinsey, Helen Light, Pennine MSK Partnership Ltd; Sharon Kerrison, Lara Smith, Jayne Budd (LCs), Stockport NHS Foundation Trust.

WALES: Laura Ingham (LC), Christine Samuel, Susan Pearson, Hayley Radford, Emma Williams, Shari Parker, Swansea Bay University Health Board.

SCOTLAND: Janet Harkess (LC), NHS Fife; Justine Griffin (LC), NHS Greater Glasgow and Clyde.

NORTHERN IRELAND: Una McKenna (LC), Northern Health and Social Care Trust.

This work was part-funded by the European Alliance of Associations for Rheumatology (EULAR) grant number HPR035 with NHS service support costs secured from the NIHR Comprehensive Local Research Network.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to report.

DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to AH. All data relevant to the study are included in the article.

ETHICS STATEMENT

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the National Research Ethics Service Committee East Midlands – Leicester South (17/EM/0409: date 16/11/2017) and the University of Salford's School of Health Sciences Ethics Panel (HSR1617-89: date 22/02/2017). All participants provided informed, written consent.

ORCID

Alison Hammond https://orcid.org/0000-0002-5266-9991
Alan Tennant https://orcid.org/0000-0002-0728-2130
Angela Ching https://orcid.org/0000-0002-3765-2534
Jennifer Parker https://orcid.org/0000-0002-2235-5748
Yeliz Prior https://orcid.org/0000-0001-9831-6254
Monique A. M. Gignac https://orcid.org/0000-0003-4445-3274
Suzanne M. M. Verstappen https://orcid.org/0000-0001-6181-0646
Rachel O'Brien https://orcid.org/0000-0002-4720-1956

REFERENCES

Acquadro, C., Joyce, C. R. B., Patrick, D. L., Ware, J. E., & Wu, A. W. (2004). *Linguistic validation manual for patient-reported outcomes (PRO) instruments*. Mapi Research Trust.

Al Dhanhani, A. M., Gignac, M. A. M., Beaton, D. E., Su, J., & Fortin, P. R. (2015). Job accommodations availability and utilization among people with lupus: An examination of workplace activity limitations and work context factors. *Arthritis Care and Research*, *67*(11), 1536–1544. https://doi.org/10.1002/art.22662

Al Dhanhani, A. M., Gignac, M. A. M., Beaton, D. E., Su, J., & Fortin, P. R. (2014). Work factors are associated with workplace activity limitations in systemic lupus erythematosus. *Rheumatology*, *53*(11), 2044–2052. https://doi.org/10.1093/rheumatology/keu242

Andrich, D., Sheridan, B. S., & Luo, G. (2015). *RUMM2030: An MS Windows computer program for the analysis of data according to Rasch unidimensional models for measurement*. RUMM Laboratory.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2007). *Recommendations for the cross-cultural adaptation of the DASH & QuickDASH outcome measures*. Institute of Work and Health. https://dash.iwh.on.ca/sites/dash/files/downloads/cross_cultural_adaptation_2007.pdf

Bellamy, N., Buchanan, W. W., Goldsmith, C. H., Campbell, J., & Stitt, L. W. (1988). Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology*, *15*, 1833–1840.

Bennett, R. M., Friend, R., Jones, K. D., Ward, R., Han, B. K., & Ross, R. L. (2009). The revised fibromyalgia impact questionnaire (FIQR): Validation and psychometric properties. *Arthritis Research and Therapy*, *11*(4), R120. https://doi.org/10.1186/ar2783

Boonen, A., Webers, C., Butink, M., Barten, B., Betteridge, N., Black, C., Bremander, A., Boteva, B., Brzezińska, O., Chauhan, L., Copsey, S., Guimarães, V., Gignac, M., Glaysher, J., Green, F., Hoving, J. L., Marques, M. L., Smucrova, H., Stamm, T. A., … Verstappen, S. M. M. (2023). EULAR points to consider supporting people with rheumatic and musculoskeletal diseases to participate in healthy and sustainable paid work. *Annals of the Rheumatic Diseases*, *82*(1), 57–64. https://doi.org/10.1136/ard-2022-222678

Brown, T., Hammond, A., Ching, A., & Parker, J. (2023). Work limitations and associated factors in working people with rheumatoid arthritis, axial spondyloarthritis, osteoarthritis or fibromyalgia. *Musculoskeletal Care*. (online first 28 March 2023). https://doi.org/10.1002/msc.1760

Calin, A., Garrett, S., Whitelock, H., Kennedy, L. G., O'Hea, J., Mallorie, P., & Jenkinson, T. (1994). A new approach to defining functional ability in ankylosing spondylitis: The development of the Bath Ankylosing Spondylitis Functional Index (BASFI). *Journal of Rheumatology*, *21*, 2281–2285.

Cichetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardised assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

Donoghue, D., PROP group, & Stokes, E. (2009). How much change is true change? The minimum detectable change of the Berg balance scale

in elderly people. *Journal of Rehabilitation Medicine*, *41*(5), 343–346. https://doi.org/10.2340/16501977-0337

Evans, J. D. (1996). *Straightforward statistics for the behavioural sciences*. Brooks/Cole Publishing.

Evers, A. W. M., Verhoeven, E. W. M., van Middendorp, H., Sweep, F. C. G. J., Kraaimaat, F. W., Donders, A. R. T., Eijsbouts, A. E., van Laarhoven, A. I. M., de Brouwer, S. J. M., Wirken, L., Radstake, T. R. D. J., & van Riel, P. L. C. M. (2014). Does stress affect the joints? Daily stressors, stress vulnerability, immune and HPA axis activity, and short-term disease and symptom fluctuations in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, *73*(9), 1683–1688. https://doi.org/10.1136/annrheumdis-2012-203143

Fisher-McAulley, G., Stanton, J., Jolton, J., & Gavin, J. (2003). Modelling the relationship between work life balance and organisational outcomes. In *Paper presented at the annual conference of the society for industrial-organisational psychology, Orlando, April 12, 2003* (pp. 1–26).

Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guidelines for studies on measurement properties of patient reported outcome measures. https://www.cosmin.nl/wp-content/uploads/COSMIN-reporting-guideline_1.pdf

Garrett, S., Jenkinson, T., Kennedy, L. J., Whitelock, H., Gaisford, P., & Calin, A. (1994). A new approach to defining disease status in ankylosing spondylitis: The Bath ankylosing spondylitis disease activity Index (BASDAI). *Journal of Rheumatology*, *21*, 2286–2291.

Gignac, M. A. M., Backman, C. L., Kaptein, S., Lacaille, D., Beaton, D. E., Hofstetter, C., & Badley, E. M. (2012). Tension at the borders: Perceptions of role overload, conflict, strain and facilitation in work, family and health roles among employed individuals with arthritis. *Rheumatology*, *51*(2), 324–332. https://doi.org/10.1093/rheumatology/ker317

Gignac, M. A. M., & Cao, X. (2009). "Should I tell my employer and co-workers I have arthritis?" A longitudinal examination of self-disclosure in the workplace. *Arthritis Care and Research*, *61*(12), 1753–1761. https://doi.org/10.1002/art.24889

Gignac, M. A. M., Cao, X., Lacaile, D., Anis, A. H., & Badley, E. M. (2008). Arthritis-related work transitions: A prospective analysis of reported productivity losses, work changes, and leaving the labour force. *Arthritis Care and Research*, *59*(12), 1805–1813. https://doi.org/10.1002/art.24085

Gignac, M. A. M., Lacaille, D., Beaton, D. E., Backman, C. L., Cao, X., & Badley, E. M. (2014). Striking a balance: Work-health- personal life conflict in women and men with arthritis and its association with work outcomes. *Journal of Occupational Rehabilitation*, *24*(3), 573–584. https://doi.org/10.1007/s10926-013-9490-5

Gignac, M. A. M., Sutton, D., & Badley, E. M. (2006). Re-examining the arthritis-employment interface: Perceptions of arthritis-work spillover among employed adults. *Arthritis Care and Research*, *55*(2), 233–240. https://doi.org/10.1002/art.21848

Gignac, M. A. M., Sutton, D., & Badley, E. M. (2007). Arthritis symptoms, the work environment, and the future: Measuring perceived job strain among employed persons with arthritis. *Arthritis Care and Research*, *57*(5), 738–747. https://doi.org/10.1002/art.22788

Gilworth, G., Chamberlain, A., Harvey, A., Woodhouse, A., Smith, J., Smith, G., & Tennant, A. (2003). Development of a work instability scale for rheumatoid arthritis. *Arthritis and Rheumatism*, *49*(3), 349–354. https://doi.org/10.1002/art.11114

Gilworth, G., Emery, P., Barkham, N., Smyth, N. G., Helliwell, P., & Tennant, A. (2009). Reducing work disability in Ankylosing Spondylitis – Development of a work instability scale for AS. *BMC Musculoskeletal Disorders*, *10*(1), 68. https://doi.org/10.1186/1471-2474-10-68

Gossec, L., Paternotte, S., Aanerud, G. J., Balanescu, A., Boumpas, D. T., Carmona, L., de Wit, M., Dijkmans, B. A. C., Dougados, M., Englbrecht, M., Gogus, F., Heiberg, T., Hernandez, C., Kirwan, J. R., Mola, E. M., Cerinic, M. M., Otsa, K., Schett, G., Scholte-Voshaar, M., … Kvien, T. K. (2011). Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient derived composite measure of impact of rheumatoid arthritis: A EULAR initiative. *Annals of the Rheumatic Diseases*, *70*(6), 935–942. https://doi.org/10.1136/ard.2010.142901

Hammond, A., Tennant, A., Ching, A., Parker, J., Prior, Y., Gignac, M. A. M., Verstappen, S., & O'Brien, R. (2023). Psychometric testing of the British-English Workplace Activity Limitations Scale in four rheumatic and musculoskeletal conditions. *Rheumatology Advances in Practice*, *7*(1), rkad028. https://doi.org/10.1093/rap/rkad028

Hayman, J. (2005). Psychometric assessment of an instrument designed to measure work life balance. *Research and Practice in Human Resource Management*, *13*, 85–91.

Health Education England. *Health literacy "how to" guide*. Health Education England. https://library.nhs.uk/wp-content/uploads/sites/4/2020/08/Health-literacy-how-to-guide.pdf

Heerkens, J. F., de Brouwer, C. P. M., Engels, J. A., van der Gulden, J. W. J., & Kant, I. (2017). Elaboration of the contextual factors of the ICF for occupational health care. *Work*, *57*(2), 187–204. https://doi.org/10.3233/WOR-172546

IBM Corp. (2019). *IBM SPSS statistics for Windows, version 26.0*. IBM Corp. Released.

Job Content Questionnaire Center. (2021). *The JCQ and JCQ2.0*. Job Content Questionnaire Centre.

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., & Amick, B. (1998). The job content questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology*, *3*(4), 322–355. https://doi.org/10.1037//1076-8998.3.4.322

Kirwan, J. R., & Reeback, J. S. (1986). Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. *British Journal of Rheumatology*, *25*(2), 26–29. https://doi.org/10.1093/rheumatology/25.2.206

Lerner, D., Amick, B. C., Rogers, W. H., Malspeis, S., Bungay, K., & Cynn, D. (2001). The work limitations questionnaire. *Medical Care*, *39*(1), 72–85. https://doi.org/10.1097/00005650-200101000-00009

Liu, Y.-Z., Wang, Y.-X., & Jiang, C.-L. (2017). Inflammation: The common pathway of stress-related diseases. *Frontiers in Human Neurosciences*, *11*, 316. https://doi.org/10.3389/fnhum.2017.00316

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

National Literacy Trust. (2017). What do adult literacy levels mean? https://literacytrust.org.uk/parents-and-families/adult-literacy/what-do-adult-literacy-levels-mean/

Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.

Office for National Statistics. (2016). Standard occupational classification SOC 2010. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.

Stratford, P. W. (2004). Getting more from the literature: Estimating standard error of measurement from reliability studies. *Physiotherapy Canada*, *56*(01), 27–30. https://doi.org/10.2310/6640.2004.15377

Tang, K., Beaton, D. E., Lacaille, D., Gignac, M. A. M., Zhang, W., Anis, A. H., & Bombardier, C. (2010). The work instability scale for rheumatoid arthritis (RA-WIS): Does it work in osteoarthritis? *Quality of Life Research*, *19*(7), 1057–1068. https://doi.org/10.1007/s11136-010-9656-y

Tang, K., Escorpizo, R., Beaton, D. E., Bombardier, C., Lacaille, D., Zhang, W., Anis, A. H., Boonen, A., Verstappen, S. M., Buchbinder, R.,

Osborne, R. H., Fautrel, B., Gignac, M. A., & Tugwell, P. S. (2011). Measuring the impact of arthritis on worker productivity: Perspectives, methodological issues, and contextual factors. *Journal of Rheumatology*, *38*(8), 1776–1790. https://doi.org/10.3899/jrheum.110405

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis and Rheumatism*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108

Tennant, A., Hillman, M., Fear, J., Pickering, A., & Chamberlain, M. A. (1996). Are we making the most of the Stanfird Health Assessment Questionnaire? *British Journal of Rheumatology*, *35*(6), 574–578. https://doi.org/10.1093/rheumatology/35.6.574

Teresi, J. A., Kleinman, M., & Ocepek-Welikson, K. (2000). Modern psychometric methods for detection of differential item functioning: Application to cognitive assessment measures. *Statistics in Medicine*, *19*(11–12), 1651–1683. https://doi.org/10.1002/(sici)1097-0258(20000615/30)19:11/12<1651::aid-sim453>3.0.co;2-h

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were prosed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Wolfe, F. (2001). Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): Analyses in 2491 rheumatoid arthritis patients following leflunomide initiation. *Journal of Rheumatology*, *28*, 982–989.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hammond, A., Tennant, A., Ching, A., Parker, J., Prior, Y., Gignac, M. A. M., Verstappen, S. M. M., & O'Brien, R. (2023). Psychometric testing of the British-English Long-Term Conditions Job Strain Scale, Long-Term Conditions Work Spillover Scale and Work-Health-Personal Life Perceptions Scale in four rheumatic and musculoskeletal conditions. *Musculoskeletal Care*, 1–16. https://doi.org/10.1002/msc.1774