Linguistic validation, validity, and reliability of the British English version of the Workplace Activity Limitations Scale in employed people with rheumatoid arthritis, ankylosing spondyloarthritis, osteoarthritis and fibromyalgia .

Alison Hammond (A.Hammond@salford.ac.uk)
University of Salford https://orcid.org/0000-0002-5266-9991
Alan Tennant
University of Leeds Faculty of Medicine and Health
Angela Ching
University of Salford School of Health and Society
Jennifer Parker
University of Salford School of Health and Society
Yeliz Prior
University of Salford School of Health and Society
Monique Gignac
Institute for Work and Health
Suzanne Verstappen
University of Manchester School of Medicine: The University of Manchester Faculty of Biology Medicine and Health
Rachel O'Brien
Sheffield Hallam University College of Health Wellbeing and Life Sciences

Research Article

Keywords: patient reported outcomes, work, work rehabilitation, arthritis, musculoskeletal, rehabilitation

Posted Date: August 30th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1744048/v1

Abstract

Purpose: The Workplace Activity Limitations Scale (WALS) is a patient reported outcome measure of work ability developed in Canada for use in arthritis. The aims of this study were to linguistically validate a British English version of the WALS, and then psychometrically test this in rheumatoid arthritis (RA), ankylosing spondyloarthritis (AS), osteoarthritis (OA) and fibromyalgia (FM).

Methods: Phase 1: The WALS was forward translated, reviewed by an expert panel and cognitive debriefing interviews conducted. Phase 2: Participants completed a postal questionnaire booklet. Construct validity was examined by fit to the Rasch measurement model. Concurrent validity included testing between the WALS and the Work Limitations Questionnaire-25 (WLQ-25). Two weeks later, participants were mailed a second questionnaire booklet for test-retest reliability.

Results: Phase 1: Minor wording changes were made. Phase 2: 831 employed participants completed questionnaires: 267 men; 564 women; 53.5 (SD 8.9) years of age; with condition duration 7.7 (SD 8.0) years. The WALS satisfied Rasch model requirements when implemented in a bi-actor equivalent solution. A WALS Rasch transformation table was created. Concurrent validity was strong with the: WLQ-25 (RA r_s = 0.78; AS r_s = 0.83; OA r_s = 0.63; FM r_s = 0.64). Internal consistency was consistent with group use (α = 0.80 to 0.87). Test-retest reliability was excellent (ICC (2,1) at 0.90 or above.

Conclusions: A reliable, valid British English version of the WALS is now available for use in the United Kingdom.

Background

Work participation (i.e., paid work) is important for people with rheumatic and musculoskeletal disorders' (RMD) health and well-being and for societies to prosper. A recent study in the United Kingdom (UK) found that healthy working life expectancy from age 50 onwards was a third lower for people with osteoarthritis (OA) compared to those without OA (i.e., 5.68 years (95% CI: 5.29, 6.07) compared to 10.00 years (95% CI: 9.74, 10.26)) [1]. People with RMD are less likely to be employed compared to those without a long-term health condition (62% versus 81%) [2].

Working people with RMD may struggle to manage at work, with their health condition causing at-work productivity loss (i.e., reduced work output due to health problems; or presenteeism) [3]. This is an important outcome to target for improvement in medical, rehabilitation and vocational interventions, as well as from people with RMD' perspectives [4]. Accordingly, there is a need for measures that focus on specific types of limitations at work to direct such interventions. Such measures should be tested across a range of RMD.

The Outcome Measures in Rheumatology (OMERACT) Work Productivity Group identified two multi-item patient reported outcome measures (PROM) of at-work productivity loss suitable for use in RMD [5, 6]. the Work Limitations Questionnaire-25 (WLQ-25), measuring the amount of time a person experiences difficulties with work activities [7]; and the Workplace Activity Limitations Scale (WALS), measuring the degree of difficulty with work activities [8]. The WALS is a measure of work ability and WLQ-25 of work productivity [4]. In a previous study, the WALS was marginally considered preferable to the WLQ-25 as an outcome measure by people with rheumatoid arthritis (RA) and OA [9].

The WALS was developed in Canada and has been used in Canadian studies with people with inflammatory arthritis (IA: i.e., RA, psoriatic arthritis, ankylosing spondyloarthritis (AS)), OA, lupus and scleroderma [10–15]. Psychometric testing has primarily been conducted in Canada. Content validity was evaluated in RA or OA (n = 250), and considered: comprehensive; with good comprehensibility, length, response options; and the number of "not applicable" items low, indicating good content relevance [9], with low respondent burden [16]. There is no evidence for internal construct validity (unidimensionality). There is good support for construct validity in RA and OA with other at-work productivity measures (RA-Work Instability Scale (RA-WIS) r = 0.77; WLQ-25 Index r = 0.61 [17]; WLQ-PDmod (i.e., the physical demands modified sub-scale) r = 0.79, Work Productivity Activity Impairment (WPAI) scale r = 0.81 [6]; and health measures (Health Assessment Questionnaire (HAQ) r = 0.66; pain r = 0.64; arthritis severity r = 0.60) [17]. Reliability (internal consistency: Cronbach's alpha) ranges from 0.78 to 0.87 in IA, OA and lupus [13, 16, 17]. Test-retest reliability in "workers with arthritis" (n = 37) reporting no change in health over a two-week period, was deemed good (Intra-Class Correlation Coefficient (ICC (2,1) = 0.93) [5]. The Minimum Detectable Change (MDC_{95}) was reported as 5.95, and Minimal (clinically) Important Difference (MID) -2.4 to -4.9 (improvement) and 1.0 to 1.1 (deterioration), dependent on the anchor used, in "workers with arthritis" [6].

The WALS was developed in Canadian English. Outcome measures should be linguistically validated (i.e., translated and culturally adapted) into the language of the target country, even if a different form of the same language, such as English. Although most of the language is understandable, some words have different meanings in British English, e.g., "subway"refers to rapid transport systems in North America but an underpass for crossing roads in the UK. Accordingly, the WALS needs linguistic validation into British English and then psychometric testing for use in the UK [18].

The aims of this study were to: linguistically validate, investigate content validity, and evaluate the psychometrics of a British English WALS amongst employed people with RA. AS, lower limb OA and fibromyalgia (FM) in the UK. Psychometric testing of measures should include a combination of classical testing and item response theory (e.g., Rasch analysis) to establish psychometric properties, including unidimensionality [19].

Methods

Ethical approval

was obtained from the National Research Ethics Service Committee East Midlands – Leicester South (17/EM/0409) and the University of Salford's School of Health Sciences Ethics Panel. All participants provided written, informed consent.

Study Design

Cross-cultural adaptation, followed by cross-sectional surveys to establish psychometric properties of the WALS. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklists for assessing methodological quality and reporting guidelines were followed [19, 20].

Participants and recruitment

Patients were identified by research facilitators or therapists in 47 UK National Health Service (NHS) Trusts (41 secondary care and six community Rheumatology, Orthopaedic or Therapy out-patient clinics). We also recruited some possible participants from our research group's Arthritis Volunteer Register. Participants were eligible if: at least 18 years of age; in paid employment for at least one day a week (including self-employed); currently at work (or if on short-term sick leave, i.e., less than four weeks, participation was delayed until at work); and a confirmed primary diagnosis of: RA or undifferentiated inflammatory arthritis (UIA); AS or axial spondylopathy (AxSpA), OA (knee and/or hip); or FM. Diagnoses were confirmed by a Rheumatologist for RA/IA and AS/AxSpA; or a Rheumatologist, Orthopaedic Surgeon, General Practitioner, or extended scope practitioner physiotherapist for OA and FM. There were no restrictions on RMD duration. Participants needed to be able to read, write and understand British English. Exclusion criteria were: on long-term sick leave (as unable to complete the work measures); and unable to provide informed consent. Patients were identified using these criteria, given a short study explanation and information pack (introductory letter, participant information sheet, reply form, and Freepost envelope to the research team). The reply form included diagnosis, employment, and sick leave status, to further check eligibility criteria. The patient could return the reply form themselves or provide written agreement that NHS staff could do so on their behalf.

For those contacted for Phase 1, a study explanation was provided by telephone and written consent given prior to interview. For those in Phase 2, the questionnaire booklet included a consent form on the front.

Phase 1: Linguistic validation, cross-cultural adaptation, and content validity

The following procedures were used [21]:

Forward translation: two translators (a rheumatology researcher familiar with the WALS (AH), and a non-health professional (experienced teacher, including of English: JG) unfamiliar with the WALS) independently reviewed the WALS to identify words requiring changing into British English and use of Plain English (i.e., simplifying words and phrases).

Translation synthesis

the two translators discussed and agreed recommended changes.

Backward translation

was not required as the translation was into another form of English.

Expert committee review: The committee included: one translator (AH); three occupational therapists experienced in work and musculoskeletal conditions (YH, TW, RO'B); the WALS developer MG: Canadian-English speaker); experienced PROMS researchers (AT, AH, YP, SV) and two patient research partners (AP, SK). The committee discussed the synthesised translation, made additional recommendations, and agreed and approved the draft British English WALS. This process ensures semantic, idiomatic, experiential, and conceptual equivalence.

Field testing of the draft WALS and content validity

Cognitive debriefing interviews were used to investigate the WALS from people with RMD' perspectives [18]. PROM content validity should be assessed by experts, i.e., patient/ public representatives of the target populations [22]. At least 10 in each target group should be included [23]. Participants were mailed a paper questionnaire booklet, including the draft British English WALS, to complete at home, and asked to consider WALS ease of completion, item relevance and if anything important was missing. Within two weeks, they were interviewed, face-to-face or by telephone, about comprehensiveness (1 = not relevant; 5 = extremely relevant; and any missing items) and comprehensibility (instructions, content, layout). Findings were discussed with the expert committee, further changes made and the final British English WALS agreed.

Content validity was further examined by linking the WALS to the Activities and Participation component of the International Classification of Functioning, Disability and Health (ICF) Core Set for Vocational Rehabilitation [24, 25]. The Flesch-Kincaid Grade score was calculated using Microsoft Word to check readability was similar to the original WALS [26].

Phase 2: Psychometric testing

Data collection

Participants were mailed a paper questionnaire booklet to complete at home (Test 1: T1). Two weeks after return, they were mailed a second paper questionnaire (Test 2: T2), to assess test-retest reliability. If either were not returned, at two weeks participants were sent a reminder letter, followed at four weeks by a further letter and copy of the questionnaire booklet. The T1 booklet included items on demographic, disease and employment characteristics: age, gender, living arrangements, education status, condition duration, medication regimen, employment status and job title, to allow coding to job skill level (1 = elementary occupations; 2 = requiring compulsory education and work-related training; 3 = post-compulsory education (sub-degree) or longer work experience; 4 = degree level education or equivalent experience [27]). Data were collected as part of a wider study testing six other contextual factor work-related measures. At T2, participants did not have knowledge of previous scores.

Instruments

To test construct (concurrent) validity, at T1 we included the following work and health measures. For all, a higher score indicates worse status.

Work measures
British English WALS: 12 items, measured on a 0–3 scale of difficulty performing work tasks (0 = no difficulty; to 3 = unable to do (Supplementary File 1). WALS content is specific to arthritis with items created through literature review [10]. It includes: eight physical activity items (e.g., working with hands, standing, moving around inside, commuting); three about managing work (i.e., work hours, pace and job demands); and one mental demand (concentration at work) [12]. Instructions state to answer without help from others, use of special gadgets or equipment, so as not confounded by use of workplace behavioural coping strategies [10]. Recall period is not specified. Items recorded as "not applicable to my job" are assigned a score of 0. The scoring allows up to three missing items, which can be imputed using individual's mean or median scores (depending on data distribution). A total summed score is calculated (0–36). A WALS score of 0–4 is considered to indicate a low level of work limitations, 5–8 = moderate; and ≥ 9 = high [13]. A score ≥ 9 is associated with greater need for work accommodations, absenteeism and job disruptions, compared to those scoring < 5 [13].

WLQ-25: a reliable, valid measure including 25 items in four sub-scales (1–5 scale), indicating the percentage time in the past two weeks a person was limited in physical work demands, time demands, mental-interpersonal demands and output demands [7]. From these, the WLQ Percentage Productivity Loss [7] and Summed scores [28] can be created.

WIS: measured using the RA-WIS in RA, OA, and FM and AS-WIS in AS [29–31]. This measures the degree of mismatch between work abilities and job demands. There is evidence for reliability and validity for the RA-WIS in RA and OA, but not yet for FM; and for the AS-WIS. The RA-WIS includes 23 true/false items and the AS-WIS 20 items. Both measures have cut-points indicating low, moderate and high work instability (Table 2).

WPAI (General Health)

a reliable, valid measure of six items from which Percentage Overall Work Impairment due to Health (in past seven days) is calculated [32].

Health status measures

Perceived health status: measured using a 5-point Likert scale "Considering all the ways that your condition affects you, how have you been over the past month?" (1 = very good (no symptoms; no limitation of normal daily activities); to 5 = very poor (very severe, intolerable symptoms; unable to do many normal daily activities)).

Perceived change in health status: At T2 only, measured using a 5-point Likert scale "Overall, how much is your arthritis/ condition troubling you now compared to when you last completed this questionnaire?" (1 = much less; 3 = about the same; 5 = much more).

Condition-specific health measures:

Four condition specific questionnaire booklets were used. Participants completed only those measures relevant to their condition.

RA

Rheumatoid Arthritis Impact of Disease (RAID): includes seven 0–10 numeric rating scales (NRS): pain, fatigue, sleep, functional disability, coping, physical and emotional well-being. A total score is created from the sum of weighted NRS scores [33].

HAQ: physical function evaluated by 20 daily activities rated on a 0–3 scale (0 = not at all difficult; 3 = unable to do) [34]; scored using the HAQ20 method, i.e., all 20 items are summed (0–20 = mild; 21–40 = moderate; 41–60 = severe disability) without adjustment for using aids and devices [35].

AS

Bath Ankylosing Spondyloarthritis Disease Activity Index (BASDAI): six 10cm. visual analogue scales (VAS) of symptom severity (fatigue, spinal pain, other joint pain/swelling, localised tenderness, morning stiffness, duration of morning stiffness), from which an average score (0–10) is calculated. Scores > 4 indicate active disease [36].

Bath Ankylosing Spondyloarthritis Functional Index (BASFI): Ten 10cm. VAS of physical function (mobility), from which an average score (0–10) is calculated [37].

OA

Western Ontario McMaster Universities Osteoarthritis Index (WOMAC): two of the three sub-scales: pain (five items); and physical function (17 items), scored on 0–4 scales (0 = none; 5 = extreme), from which total scores for each sub-scale are calculated [38].

FM

Revised Fibromyalgia Impact Questionnaire (FIQR): three sub-scales rated on 0–10 NRS: overall impact (two items); symptoms (10 items); and function (nine items). Sub-scale and overall total scores are calculated [39].

Sample size

As Rasch analysis was used to assess internal construct validity (unidimensionality), enough cases are needed within each condition group to test for invariance across groups [40]. The sample does not need to be representative, as the Rasch model is independent of distribution, but should have a good distribution across the work activity limitation domain. A minimum of 150 responses is required for Rasch analysis, although we aimed to collect up to 250 to ensure a broad spread of responses. At least 79 sets of repeated responses were required 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 correlation of 0.7 is deemed a minimum acceptable level [41].

Statistical analyses

Demographic, work, and disease measures were summarised descriptively, as appropriate. RUMM 2030 + software was used for Rasch analysis [42]. As all work and health measures either consisted of ordinal data, or were not normally distributed, non-parametric statistical tests were conducted using the Statistical Package for the Social Sciences (SPSS) v26 [43].

Compliance (missing data): the number (%) of missing data items and WALS which could not be scored were identified.

Internal construct validity: The primary analytical strategy was testing the fit of the WALS for each condition to the Rasch Measurement Model to determine reliability and internal construct validity [44]. Given the requirements for fit, a hierarchical strategy was used (Supplementary Table 1). With level 1 as the priority (individual item fit), all requirements for model fit must be met. Should a Level 5 solution be unavailable (bi-factor solution on alternative items), item deletion will be considered (Level 6). If this fails then Level 7 will be used to test for a valid ordinal scale, and if this fails then Level 8 indicates no valid ordinal scale. Details of the Rasch analysis undertaken are in Supplementary File 2 and described elsewhere [45].

Construct (concurrent) validity: was assessed using Spearman's correlations with work and health measures. Correlations were deemed: 0.8-1 = very strong; 0.6–0.79 = strong; 0.4–0.59 = moderate; 0.20–0.39 = weak; and 0-0.19 = very weak [46]. We hypothesised that, in the four condition groups, there would be: moderate to strong correlations between the WALS and scores for the three work measures: WLQ-25 (Percentage Productivity Loss and Summed scores), WIS (RA- and AS-WIS) and WPAI; and moderate correlations with severity of perceived health status, and condition-specific symptoms and physical function scales.

Discriminant validity: was assessed using Kruskal-Wallis tests to evaluate differences in perceived health status between participants reporting very poor/poor; fair; good/very good health status. A p-value of $p \leq 0.05$ was considered significant.

Internal consistency: was assessed using Cronbach's alpha. Results of ≥ 0.8 were deemed good to excellent: ≥ 0.9 is consistent with individual use; and > 0.7 with group-level use [46].

Test-retest reliability was assessed in those reporting their health was "the same" at T2, using Spearman's correlations and ICC (2,1): two-way random consistency, average measures model. An ICC ≥ 0.75 was considered excellent and 0.5–0.74 moderate [47]. Reliability of individual WALS items was calculated using linear weighted kappa, with levels of agreement as < 0.20 = poor; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = good; 0.81-1.00 = very good [46].

Sensitivity to change: was assessed by calculating Standard Error of Measurement (SEM) and the Minimal Detectable Change$_{95}$ (MDC$_{95}$) scores, i.e. a statistical estimate of the smallest detectable change corresponding to change in ability [47, 48]. The formulae used were: SEM = $s\sqrt{(1 - r)}$, where s = the mean plus standard deviation (SD) of T1 and T2 difference, r = the reliability coefficient for the test, i.e. Pearson's correlation co-efficient between T1 and T2 values. Thereafter the MDC$_{95}$ was calculated using the formula: MDC$_{95}$ = SEM × $\sqrt{2}$ × 1.96 [48].

Floor and ceiling effects: were considered present if > 15% of participants achieved either the lowest or highest scores in the WALS [50].

Results

Phase 1: Linguistic validation, cross-cultural adaptation, and content validity

Following forward translations and synthesis, the expert panel reviewed these and agreed the following changes: to have a root question "how much difficulty do you have…", rather than each item starting with this; item 1 (travel) changing "subway" to "train" and including active travel (walking, cycling); item 2 (get around the workplace) changing "hallways" to "corridors" and adding "machinery" to include physical work settings; item 6 (work with hands) widening examples to include keyboard/touchscreen (as most jobs now involve computer usage), tools and operating machinery (to be inclusive of physical work) and adding smartphone to "hold a phone"; item 9 (managing hours of work) changing "schedule" to "shifts," as more commonly used in British English; and item 12 (concentration) changing "due to" arthritis" to "condition", as those with FM may not consider their diagnosis as arthritis.

Cognitive debriefing interviews were conducted with 48 participants (face-to-face n = 6; telephone n = 42) (Table 1), with results reviewed by the expert panel to determine any further changes in the WALS required. Most participants considered the WALS comprehensive, with items very or extremely relevant for their condition, with no significant differences between groups (Supplementary Table 2). Only six suggested additional items, although only by one each and so not included. These were: "driving for work"; "parking near work"; "opening door handles, bottles and jars"; "going to the [work] bathroom (turning door handles and taps)"; "dealing with co-workers and the public" and "having to explain myself [condition] to people." Most (43/48) considered the WALS comprehensible, with instructions, content, and layout easy to understand. Only five stated these were "partly easy." Of these, one recommended moving the "not applicable" column

to the first response option, rather than last, which was changed. As only one person each raised the following issues, no changes were made: one participant (AS) misunderstood instructions, indicating all activities were “not applicable” as able to do them despite pain; and another participant (FM) reported taking several attempts to assimilate longer items (e.g., item 6). Three noted problems with the response options, with one each stating: the gap between “some” and “a lot” is too large and an intermediate option needed; frequency of difficulty rather than amount would be better; and preferring a focus on ability rather than difficulty. No problems were reported with the lack of time frame in the instructions.

Table 1
Phase 1 and 2 demographic data: RA, AS, OA and FM

	RA		AS		OA		FM	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
n =	12	294	10	199	13	173	13	156
Sex: M:F n (%)	5:7	76 (26): 218 (74)	4:6	124 (62.30): 75 (37.7)	4:9	54 (31): 119 (69)	2:11	10 (6): 146 (94)
Age (years): - mean (SD)	57.33 (6.77)	53.47 (8.97)	33 (14.62)	46.96 (10.24)	55.92 (6.70)	56.49 (7.21)	39.69 (9.11)	45.71 (10.05)
Job skill level: n (%)	3	149 (51)	5	66 (33.1)	8	84 (48.60)	7	95 (61)
- 1 and 2	9		5	133 (66.9)	5		6	61 (39)
- 3 and 4	-	142 (48)	-	-	-	88 (50.80)	-	-
- Missing		3 (1)				1 (0.70)		
Disease duration (years): mean (SD)	18.08 (11.93)	7.66 (7.97)	12.70 (9.78)	12.33 (10.40)	12.35 (10.60)	5.28 (8.00)	5.38 (3.55)	2.99 (4.17)
Phase 2 only								
Symptom duration (years) mean (SD)	9.33 (8.52)		18.97 (11.75)		8.10 (8.84)		8.36 (7.16)	
Living conditions: n(%) - With spouse/family/significant other	241 (82)		179 (89.9)		143 (83)		139 (89)	
Children < 18y living at home n (%)	69 (23)		68 (34.2)		31 (18)		56 (36)	
Educational Level (ISCED): n (%)	27 (9.2)		14 (7)		17 (10)		7 (4)	
- No formal qualifications	148 (50.3)		100 (50.3)		91 (53)		76 (49)	
- Secondary/non tertiary	117 (39.8)		84 (42.2)		61 (35)		73 (47)	
- Tertiary	2 (0.7)		1 (0.5)		4 (2))		-	
- Missing								
Full-: part-time work n (%)	160 (54.4): 134 (45.6)		150 (75.4): 49 (24.6)		106 (61.30): 67 (38.7)		70 (45):86 (55)	
Hours worked: mean (SD)	33.24 (12.47)		37.77 (10.44)		34.16 (11.66)		31.50 (10.56)	
Self-employed n (%)	63 (21.4)		34 (17.1)		21 (12.10)		18 (11.5)	
Physical demands of job: n (%)	101 (34.4)		83 (41.7)		53 (30.70)		61 (39.1)	
- None/a little	37 (12.6)		175 (8.9)		22 (12.70)		14 (9)	
- Noticeable	156 (53)		99 (49.8)		98 (56.60)		81 (51.9)	
- A lot/ great deal								
Medication regimen: n (%)	2 (0.7)		19 (9.5)		33 (19)		23 (15)	
- None	11 (3.7)		4 (2)		118 (69)		14 (9)	
- NSAIDS +/- analgesics	6 (2)		51 (25.6)		10 (6)		6 (4)	
- Steroids +/ - NSAIDS	103 (35)		10 (5)		-		-	
- Single DMARD	97 (33)		2 (1)		-		-	
- Combination DMARD	66 (22.4)		112 (56.3)		12 (7)		-	
- Biologic/biosimilar	-						99 (64)	
- Neuropathic analgesics (e.g., gabapentin/ pregabalin)							12 (8)	
- FM: Opiate medication								

Key: RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia; SD = standard deviation; ISCED = International Standard Classification of Education; NSAID = non-steroidal anti-inflammatory drugs; DMARD = disease modifying anti-rheumatic drugs.

The WALS was linked to 16 items in the ICF Core Set for Vocational Rehabilitation, indicating reasonable coverage, although it could potentially be linked to a further 14 items dependent on how a person interprets item 11 (managing job demands) in relation to their job (Supplementary Table 3). The Flesch-Kincaid Grade Level score was 7.6, similar to the original WALS at 7.1, indicating a reading age of 11- to 13-year-olds [26].

Phase 2: Psychometric testing

Participants

Overall, 1,359 people were referred to the study, 831 returned T1 and 622 T2 questionnaire booklets (Supplementary Fig. 1). Participant characteristics are shown in Table 1 and work and health measures in Table 2. Median time between tests was 36 (IQR 28–47) days.

Table 2
Phase 2: Participants' work and health measures in RA, AS, OA and FM

	RA (n = 294)	AS (n = 199)	OA (n = 173)	FM (n = 156)
Work measures:				
WALS:	9 (5–14)	6 (3–11)	10 (6–14)	16 (12–19)
0–36: median (IQR)				
WLQ-25 (0-100; median (IQR))	30 (10–55)	25 (5–50)	30 (10–50)	60 (40–80)
- Time Management Demands	37.50 (20–58.33)	37.50 (12.50–55.31)	41.67 (25–58.33)	58.33 (43.75–73.96)
- Physical Demands	16.67 (5.55–36.11)	13.89 (2.78–30.56)	16.67 (5.56–36.11)	44.44 (27.78–61.11)
- Mental Interpersonal Demands	20 (5–44.06)	10 (0–30)	20 (5–43.75)	45 (25–65)
- Output Demands	6.92 (3.27–11.12)	5.40 (1.71–9.36)	6.65 (3.43–11.40)	13.26 (9.20–16.53)
- WLQ-25% Productivity Loss	29.38 (14.17–43.70)	22.74 (7.08–40.03)	28.61 (15.21–45.36)	51.69 (37.30–64.62)
- WLQ-25 Summed score				
WIS (0–23 RA, OA, FM; 0–20 AS: median IQR):	13 (7.75–18)	11 (4–15)	13 (8–17)	18 (15–20)
- Low work instability n (%)	95 (32.30)	99 (49.70)	59 (34.10)	6 (3,84)
- Moderate work instability n (%)	123 (41.80)	80 (40.20)	79 (45.70)	64 (41)
- High work instability n (%)	76 (25.90)	20 (10.10)	35 (20.20)	86 (55.16)
WPAI: median (IQR):	30 (10–60)	20 (0–40)	30 (10–58.11)	66.15 (50–80)
-% overall work impairment due to health				
Health measures:				
Perceived severity	3 (2–3)	2 (2–3)	3 (3–3)	4 (3–4)
health last month (1–5; median IQR); n (%):	45 (15.30)	21 (10.60)	37 (21.40)	83 (53)
- Poor/very poor	133 (45.20)	78 (39.20)	95 (54.90)	63 (41)
- Fair	116 (39.50)	100 (50.30)	41 (23.70)	10 (6)
- Good/very good				
RA:				
RAID (0–10: median (IQR)	4.84 (3.15–6.42)	-	-	-
HAQ20 (0–60: median (IQR)	9 (3–18)	-	-	-
AS:				
BASDAI (0–10: median (IQR)		3.93 (1.95–5.87)		
BASFI (0–10: median (IQR)	-	2.97 (1.40–5.35)	-	-
OA:				
WOMAC Median (IQR)	-	-	10 (7–13)	-
- Pain (0–20)	-	-	4 (3.50–6)	-
- Stiffness (0–8)	-	-	31 (21–41.50)	-
- Physical Function (0–68)				
FM:				

Key: RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia; WALS = Workplace Activity Limitations Scale; WLQ-25 = Work Limitations Questionnaire-25; WIS = Work Instability Scale; WPAI = Work Productivity Activity Impairment; RAID = Rheumatoid Arthritis Impact of Disease; HAQ = Health Assessment Questionnaire; BASDAI = Bath Ankylosing Spondylitis Disability Index; BASFI = Bath Ankylosing Spondylitis Function Index; WOMAC = Western Ontario McMaster Universities Osteoarthritis Index; FIQ-R = Fibromyalgia Impact Questionnaire – Revised; SD = standard deviation; IQR = inter-quartile range; NRS = numeric rating scale.

For all measures: higher scores indicate more work/health problems; WIS (RA-WIS used for RA, AS and FM; AS-WIS for AS). WIS cut-points: low instability: RA-WIS < 10; AS-WIS < 11; moderate instability: RA-WIS 10–17; AS-WIS 11–18; high instability: RA-WIS > 17; AS-WIS = 19–20.

	RA (n = 294)	AS (n = 199)	OA (n = 173)	FM (n = 156)
FIQ-R (normalised scores: median IQR)	-	-	-	14 (10–17)
- Overall Impact (0–20)	-	-	-	34.50 (28.13–39)
- Symptoms (0–50)	-	-	-	19.33 (14.67–22.67)
- Function (0–30)				68.33 (54.20–77.50)
- FIQR Total (0–100)				
T1 to T2:	n = 219	n = 156	n = 131	n = 116
Time between T1 and T2; days (median, IQR)	40 (34–48)	38 (29–49.25)	30 (23.75–37)	33 (26.50–45)
Self-reported effect of health condition at T2 versus T1: n (%)	36 (16.44)	24 (15.39)	16 (12.10)	14 (12.07)
- Much/ somewhat less troublesome	136 (62.10)	99 (63.47)	78 (59.10)	54 (46.55)
- The same	47 (21.46)	31 (19.88)	38 (28.80)	48 (41.38)
- Somewhat/ much more troublesome		2 (1)		
- Missing				

Key: RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia; WALS = Workplace Activity Limitations Scale; WLQ-25 = Work Limitations Questionnaire-25; WIS = Work Instability Scale; WPAI = Work Productivity Activity Impairment; RAID = Rheumatoid Arthritis Impact of Disease; HAQ = Health Assessment Questionnaire; BASDAI = Bath Ankylosing Spondylitis Disability Index; BASFI = Bath Ankylosing Spondylitis Function Index; WOMAC = Western Ontario McMaster Universities Osteoarthritis Index; FIQ-R = Fibromyalgia Impact Questionnaire – Revised; SD = standard deviation; IQR = inter-quartile range; NRS = numeric rating scale.

For all measures: higher scores indicate more work/health problems; WIS (RA-WIS used for RA, AS and FM; AS-WIS for AS). WIS cut-points: low instability: RA-WIS < 10; AS-WIS < 11; moderate instability: RA-WIS 10–17; AS-WIS 11–18; high instability: RA-WIS > 17; AS-WIS = 19–20.

Psychometric analyses

Compliance

Missing data occurred for up to three participants for each item in each condition, i.e., 0.01%. In RA, AS and OA. WALS scores could not be calculated for three participants (with five to 12 missing items each) in RA, AS and OA. Accordingly, these participants were not included in the analysis. All FM scores could be calculated. The frequency of “not applicable” (which can be re-scored as 0) and “missing” data are shown in Supplementary Table 4.

Internal construct validity

The initial fit of the WALS to the Rasch model showed multidimensionality, caused by clusters of locally independent items in both the upper and lower part of the scale. Consequently, fit of the WALS items to the Rasch model in those with RA, AS, OA and FM was at level 4 (i.e., local-dependency cluster based-parallel form: Supplementary File 1) and is shown in Table 3. The items most easily affirmed (i.e., the transition from no to some difficulty) were: ‘Lifting, carrying or moving objects’ (RA); “Crouching, bending or kneeling” (AS, OA); and “Concentrating” (FM). The items most difficult to affirm (i.e., the transition from a lot of difficulty to unable to do) was: “Working with your hands” (RA, AS, OA and FM), particularly in FM, as the transition was five logits higher than the next threshold. No DIF was observed in any condition. The items ‘Managing the pace of work that your job requires’ and ‘Meeting your current job demands’ had a residual correlation of 0.37 (RA), 0.36 (AS), 0.45 (OA) and 0.50 (FM), where values above 0.12 would be considered indicative of local item dependency. It was these types of clusters that were used to make two parallel forms in each condition, so giving adequate fit to the model. In summary the WALS satisfied the Rasch model requirements when implemented in a bi-actor equivalent solution. The amount of variance discarded was small, giving confidence that the scale was unidimensional (albeit with a slightly inflated reliability at the item level). A transformation table was created to convert WALS raw scores to interval level scores, if required (Supplementary Table 5). A Reference Metric was also created to allow test equating of raw WALS scores with raw RA- and AS-WIS scores (Supplementary Table 6). Both the latter have clinically derived cut-points (Table 2). Direct comparison with these cut points suggests that WALS scores of 7 and 14 would indicate thresholds for moderate and high work instability, respectively, for RA, OA, AS and, provisionally for FM.

Table 3
Fit of the Workplace Activity Limitations Scale to the Rasch model in RA, AS, OA and FM.

Diagnosis/ Scale	Residuals (SD)		Chi-Square		Reliability		Dimen-sionality	DIF	ECV	Latent Correlation*
	Item	Person	Value (df)	p	PSI	α	% t-tests (LCI)			
RA	0.001	0.84	28.10 (20)	0.11	0.83	0.87	2.70	None	0.97	0.93
AS	0.43	0.73	18.70(15)	0.24	0.77	0.85	3.52	None	0.93	0.90
OA	0.17	0.92	22.80(19)	0.24	0.80	0.83	1.80	None	0.96	0.92
FM	0.19	0.81	13.40 (18)	0.77	0.80	0.80	3.21	None	0.98	0.95
Across all 4 conditions	0.41	0.97	24.1 (23)	0.40	0.84	0.87	2.8	None	0.97	0.95
Ideal values	**1.0**	**1.0**		**> 0.05**	**> 0.7**	**> 0.7**	**< 0.5**		**> 0.9**	**> 0.9**

Key: RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia; SD = Standard Deviation; PSI = Person Separation Index; α = Cronbach's alpha; LCI = Lower Confidence Interval; DIF = Differential Item Functioning; ECV = Explained Common Variance; * Between parallel forms.

Concurrent validity

As hypothesised for the four conditions, the WALS positively correlated moderately to strongly with work measures (total scores: r_s: 0.51–0.84), perceived health status (r_s: 0.42–0.71) and diagnosis-specific symptoms (r_s: 0.54–0.68) and physical function measures (r_s: 0.55–0.77) (Table 4).

Table 4
Concurrent validity of the Workplace Activity Limitations Scale with work and health measures in RA, AS, OA and FM (Spearman's correlations).

WALS (0–36) correlations with:	RA (n = 294)	AS (n = 199)	OA (n = 173)	FM (n = 156)
	r_s	r_s	r_s	r_s
Work measures				
WLQ-25 (0-100)	0.70**	0.75**	0.65**	0.57**
- Time Management Demands	0.62**	0.73**	0.50**	0.39**
- Physical Demands	0.68**	0.71**	0.62**	0.58**
- Mental Interpersonal Demands	0.71**	0.71**	0.52**	0.56**
- Output Demands	0.78**	0.83**	0.63**	0.64**
- WLQ-25% Productivity Loss	0.79**	0.84**	0.67**	0.66**
- WLQ-25 Summed score				
WIS (0–23 RA, OA, FM; 0–20 AS)	0.77**	0.84**	0.73**	0.60**
WPAI (%)	0.65**	0.80**	0.68**	0.51**
- Overall work impairment due to health				
Health measures:				
Self-reported health in last month (1–5)	0.61**	0.71**	0.53**	0.42**
RA:				
- RAID (0–10: median IQR)	0.68**	-	-	-
- HAQ20 (0–60: median IQR]	0.73**	-	-	-
AS:				
- BASDAI (0–10: Mean (SD):	-	0.68**	-	-
- BASFI (0–10: mean (SD))	-	0.77**	-	-
OA:				
WOMAC Median IQR)	-	-	0.56**	-
- Pain (0–20)	-	-	0.55**	-
- Physical Function (0–68)				
FM:				
FIQR (normalised scores)	-	-	-	0.43**
- Overall Impact (0–20)	-	-	-	0.54**
- Symptoms (0–50)	-	-	-	0.55**
- Function (0–30)	-	-	-	0.61**
- FIQR Total (0–100)				

Key: ** = correlation significant at 0.01 level; RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia ; WALS = Workplace Activity Limitations Scale; WLQ-25 = Work Limitations Questionnaire-25; WIS = Work Instability Scale; WPAI = Work Productivity Activity Impairment; RAID = Rheumatoid Arthritis Impact of Disease; HAQ = Health Assessment Questionnaire; BASDAI = Bath Ankylosing Spondylitis Disability Index; BASFI = Bath Ankylosing Spondylitis Function Index; WOMAC = Western Ontario McMaster Universities Osteoarthritis Index; FIQ-R = Fibromyalgia Impact Questionnaire – Revised.

Discriminant validity

There were significant differences between the three levels of perceived disease severity for the WALS across all four conditions, with higher perceived disease severity sub-groups scoring worse (Supplementary Table 7).

Internal consistency

This was inflated by local item dependency. Cronbach's alpha values across the four conditions were good to excellent: from 0.80 (FM) to 0.87 (RA). All are consistent with group level use (Table 3).

Test-retest reliability

At T2, 356/622 (57%) reported their condition was "the same" as at T1 and included in the analysis. For all four conditions, correlations between T1 and T2 scores were strong to very strong (r_s=0.80 and above). ICC (2,1) were excellent at 0.90 and above (Table 5). Item reliability was moderate to good

(Supplementary Table 8).

Table 5
Test-retest reliability and sensitivity to change of the Workplace Activity Limitations Scale in RA, OA, AS and FM

	n for test-retest [a]	Test 1 score* (median, IQR)	Test 2 score* (median, IQR)	Spearman's* correlation r_s	ICC (2,1)* (95% CI)	SEM*	MDC_{95}*
RA	136	8 (4–12.75)	7 (4–11)	0.83**	0.92 (0.90, 0.94)	1.15	3.17
AS	98	5 (2–9)	5 (1.75–7.25)	0.80**	0.90 (0.84, 0.93)	1.67	4.82
OA	78	8 (5.75–12.25)	7 (4–12.25)	0.81**	0.90 (0.84, 0.93)	1.83	5.08
FM	54	15 (10–19.25)	14 (11–19)	0.82**	0.90 (0.83, 0.94)	1.57	4.36

Key: [a] = participants indicating condition "about the same" at T1 and T2 with WALS scores available at both time points; ** = correlation significant at 0.01 level. RA = rheumatoid arthritis; AS = axial spondyloarthritis; OA = osteoarthritis; FM = fibromyalgia

Sensitivity to change

The MDC_{95} scores ranged from 3.17 (RA) to 5.08 (OA) in those stating their health was "the same" at T2 (Table 5)

Floor and ceiling effects

Fewer than 15% scored 0 for the WALS, indicating there was no floor effect: RA = 6/294 (2%); AS = 21/199 (10.4%); OA = 3/176 (1.7%); FM = 1/156 (0.6%). There were no ceiling effects (score 36): RA = 0/294 (0%); AS = 0/199 (0%); OA = 0/176 (0%); FM = 0/156 (0%).

Discussion

A linguistically validated British English version of the WALS is now freely available for use in the UK. An online version is also available at www.mskhub.com. The WALS already has evidence from two or more studies in Canada for content and construct validity, reliability, responsiveness, feasibility and one study for interpretability [6], primarily in RA and OA. This study extends this evidence for RA and OA and additionally provides evidence in AS and FM. The British-English WALS demonstrated good psychometric properties in working people with RA, AS, OA and FM and can be used in the UK.

We ensured linguistic and cross-cultural validity of the WALS by using a standard translation process [21] and confirming the WALS developer's approval. Example activities were updated in three items: to be reflective of active travel options (item 1), which may be more common in the UK than Canada; and items 2 and 6 to increase relevancy for those in manual jobs. Participants considered the WALS comprehensive, comprehensible, and easy to complete, indicating good content validity from the patients' perspective in these four RMD, i.e., comparable to content validity findings in RA and OA in Canada [9].

To our knowledge, this is the first study examining internal construct validity of the WALS in RA, AS, OA and FM, demonstrating fit to the Rasch model and making available a Rasch transformation table from WALS raw to interval scores. As the WALS is unidimensional, summed or (Rasch) standardised scores can be used. As hypothesised, the WALS demonstrated good construct validity with other work measures, the exception being the WLQ-25 Physical Demands sub-scale in FM. Some participants can have difficulty completing the WLQ-25 Physical Demands sub-scale, as instructions are reversed compared to the other three sub-scales [6]. Potentially, more with FM experienced such difficulty as FM can lead to greater cognitive deficits than RA, AS or OA. As hypothesised, correlations with physical function, symptom and health scales were moderate in OA and FM, but generally strong (i.e., higher than expected) in RA and AS. These findings are also comparable to those in RA and OA in Canada [17]. We also demonstrated the WALS has good discriminative validity in the four RMD, which had, to the best of our knowledge, not previously been tested.

Reliability (internal consistency) was good and comparable to findings in RA and OA in Canada [6]. Results indicate that the WALS can be used for group measurement in RA, AS, OA and FM. Identifying that WALS scores of 7 and 14 equate to RA- and AS-WIS cut points for moderate and high work instability, indicate the WALS could be useful to help identify patients' work limitations and who might benefit from work rehabilitation. Test-retest reliability was good. The evidence for test-retest reliability is extended and specific MDC_{95} for each of the four RMD were provided. These had previously only been tested in a small sample of "workers with arthritis" [5, 6].

The WALS tests intrinsic work activity impairment (i.e., capacity in ICF terms), as the instructions specify reporting difficulty without help from another person or use of gadgets or equipment. It may not therefore reflect the person's actual work ability (performance in ICF terms), i.e., with ergonomic modifications, help, and / or job accommodations. Under the UK Equality Act 2010, it is the duty of an employer to provide these (termed as "reasonable adjustments") to employees with disabilities. Clinically, and in work rehabilitation studies, using a WALS omitting instructions to answer "without help or gadgets/ equipment"

could better identify if improvement occurs following work rehabilitation and putting reasonable adjustments in place. Modified instructions could focus on how people usually do these activities. Additionally, there is no time frame in the instructions. Some work measures, e.g., the WLQ-25, ask about the last two weeks. A disadvantage of a short time frame is that the measure can only be completed by people working for at least one day during that time. Those on sick, annual, or other extended leave for more than two weeks cannot complete it. Secondly, people with RMDs may experience episodic flares or worse health. A limited time frame means completion may coincide with a period of unusual ill-health or good health. Avoiding a time scale overcomes this problem, as participants might either reflect on their difficulties when last at work or estimate difficulties. This could, however, be problematic in those on long-term sick leave if they incorrectly estimate difficulties. Particularly in intervention studies, it may be better to specify a time (e.g., three months). Future research could psychometrically test a WALS with modified instructions.

A strength of this study was that we had relatively large samples of people with RA, AS, OA and FM recruited from a wide variety of NHS out-patient clinics, meaning results are representative for people accessing secondary or community care. Limitations were that fewer people with FM had stable self-reported health between T1 and T2, compared to the other conditions, resulting in a smaller sample for test-retest reliability than required. Responsiveness (i.e., longitudinal validity) still needs testing and minimal clinically important differences (MCID) also establishing in the UK. Further testing in other RMD is required.

Conclusions

Overall, psychometric testing of the British English WALS demonstrated good validity and reliability in employed people with RA, AS, OA and FM in the UK. The WALS meets most recommendations of the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklists for methodological quality and reporting [19, 20]. Accordingly, the British English WALS can be used in the UK for these four RMDs.

Declarations

Acknowledgements:

The authors wish to thank: all the study participants for their time in completing 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 Hospital's 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.Holy 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

Funding

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

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Authors Contributions

Alison Hammond, Alan Tennant and Yeliz Prior contributed to the study conception and design. All authors contributed to Phase 1 of the study. Phase 2: Material preparation and data collection were performed by Alison Hammond, Yeliz Prior, 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 written by Alison Hammond and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request, following completion of associated studies.

Ethics approval

*This study was performed in line with the principles of the Declaration of Helsinki. A*pproval 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). Approvals from the Research and Development departments were obtained at each Participant Identification Centre. All participants provided informed, written consent.

Consent to participate

Informed consent was obtained from all individual participants included in the study

Consent for publication

Not applicable

References

1. Lynch M, Bucknall M, Jagger C, Wilkie R. Healthy working life expectancy at age 50 for people with and without osteoarthritis. in local and national English populations. Nature Scientific Reports 2022. https://www.nature.com/articles/s41598-022-06490-3.pdf?origin=ppub Accessed 6 June 2022.
2. Versus Arthritis. The State of Musculoskeletal Health 2021. https://www.versusarthritis.org/about-arthritis/data-and-statistics/the-state-of-musculoskeletal-health/ Accessed 6 June 2022.
3. Ravinskaya M, Verbeek JH, Langendam M, Daams JG, Hulshof CTJ, Madan I, Verstappen SMM < Hagendijk M, Kunz R, Hoving JL. Extensive variability of work participation outcomes measured in randomized controlled trials: a systematic review. J Clin Epidemiology. 2022;142:60–99.
4. Verstappen SMM, Lacaille D, Boonen A, Escorpizo R, Hofstetter C, Bosworth A, Leong A, Leggett S, Gignac MAM, Wallman JK, Ter Wee MM, Berghea F, Agaliotis M, Tugwell P, Beaton D. Considerations for Evaluating and Recommending Worker Productivity Outcome Measures: An Update from the OMERACT Worker Productivity Group. J Rheumatol. 2019;46:1401–5.
5. Tang K, Boonen A, Verstappen SMM, Escorpizo R, Luime JJ, Lacaille D, Fautrel B, Bosworth A, Cifaldi M, Gignac MAM, Hofstetter C, Leong A, Montie P, Petersson IF, Purcaru O, Bombardier C, Tugwell PS, Beaton DE. Worker Productivity Outcome Measures: OMERACT Filter Evidence and Agenda for Future Research. J Rheumatol. 2014;41:165–76.
6. Beaton DE, Dyer S, Boonen A, Verstapapen SMM, Escorpizo R, Lacaille DV, Bosworth A, Gignac MAM, Leong A, Purcaru O, Leggett S, Hofstetter C, Petersson IF, Tang K, Fautrel B, Bombardier C, Tugwell PS. OMERACT Filter Evidence Supporting the Measurement of At-work Productivity Loss as an Outcome Measure in Rheumatology Research. J Rheumatol. 2016;43:214–22.
7. Lerner D, Amick BC, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. Med Care. 2001;39:72–85.
8. Gignac MAM, Sutton D, Badley EM. Arthritis Symptoms, the Work Environment, and the Future: Measuring Perceived Job Strain Among Employed Persons with Arthritis. Arthr Care Res. 2007;57:738–47.
9. Tang K, Beaton D, Lacaille D, Gignac MAM, Bombardier C. Canadian Arthritis Network Work Productivity Group. Sensibility of five at-work productivity measures was endorsed by patients with osteoarthritis or rheumatoid arthritis. J Clin Epidemiology 2013; 66:546–56.
10. Gignac MAM. Arthritis and Employment: An Examination of Behavioral Coping Efforts to Manage Workplace Activity Limitations. Arthr Care Res. 2005;53:328–36.
11. Gignac MAM, Cao X, Lacaille D, Anis AH, Badley EM. Arthritis-Related Work Transitions: A prospective Analysis of Reported Productivity Losses, Work Changes, and Leaving the Labor Force. Arthr Care Res. 2008;59:1805–13.
12. Gignac MAM, Cao X. "Should I Tell My Employer and Co-workers I Have Arthritis?" A Longitudinal Examination of Self-Disclosure in the WorkPlace. Arthr Care Res. 2009;61:1753–61.

13. Gignac MAM, Cao X, Tang K, Beaton DE. Examination of Arthritis-Related Work Place Activity Limitations and Intermittent Disability Over Four-and-a-Half Years and Its Relationship to Job Modifications and Outcomes. Arthr Care Res. 2011;63:953–62.
14. Al Dhanhani AM, Gignac MAM, Beaton DE, Su J, Fortin PR. Work factors are associated with workplace activity limitations in systemic lupus erythematosus. Rheumatology. 2014;53:2044–52.
15. Jetha A, Johnson SR, Gignac MAM. Unmet Workplace Support Needs and Lost Productivity of Workers With Systemic Sclerosis: A Path Analysis Study. Arthr Care Res. 2021;73:423–32.
16. Tang K, Beaton DE, Boonen A, Gignac MAM, Bombardier C. Measures of work disability and productivity: Rheumatoid Arthritis Specific Work Productivity Survey (WPS-RA), Workplace Activity Limitations Scale (WALS), Work Instability Scale for Rheumatoid Arthritis (RA-WIS), Work Limitations Questionnaire (WLQ), and Work Productivity and Activity Impairment Questionnaire (WPAI). Arthritis Care Res. 2011;63(Suppl 11):337–49.
17. Beaton DE, Tang K, Gignac MA, Lacaille D, Badley EM, Anis AH, et al. Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. Arthritis Care Res. 2010;62(1):28e37.
18. Acquadro C, Joyce CRB, Patrick DL, Ware JE, Wu AW. **Linguistic Validation Manual for Patient-Reported Outcomes** (**PRO**) **Instruments**. Mapi Research Trust, Lyon, France. 2004.
19. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19:539–49.
20. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN Reporting guidelines for studies on measurement properties of patient reported outcome measures. 2021 https://www.cosmin.nl/wp-content/uploads/COSMIN-reporting-guideline_1.pdf Accessed 6 June 2022.
21. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Recommendations for the Cross-Cultural Adaptation of the DASH & QuickDASH Outcome Measures. 2007. Toronto, ON: Institute of Work and Health. https://dash.iwh.on.ca/sites/dash/files/downloads/cross_cultural_adaptation_2007.pdf. Accessed. *6 June.2022.*.
22. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: a practical guide. Cambridge (UK): Cambridge University Press; 2011.
23. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21:651–7.
24. Finger ME, Escorpizo R, Glässel A, Gmünder HP, Lückenkemper M, Chan C, Fritz J, Studer U, Ekholm J, Kostanjsek N, Stucki G, Cieza A. ICF Core Set for vocational rehabilitation: results of an international consensus conference, Disabil Rehabil 2012:34; 429–438.
25. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learnt. J Rehabil Med. 2004;37(4):212–8.
26. JISC Tech Dis Inclusion Technology Advice. How to use the Readability statistics. 2013. https://sshs.exeter.ac.uk/media/universityofexeter/schoolofsportandhealthsciences/documents/Guidelines_for_checking_readability_of_Participant_Infor Accessed 6 June 2022.
27. Office for National Statistics. Standard Occupational Classification 2010. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010. Accessed 6 June 2022.
28. Roy JS, MacDermid JC, Amick BC, Shannon HS, McMurtry M, Roth JH, et al. Validity and responsiveness of presenteeism scales in chronic work-related upper extremity disorders. Phys Ther. 2011;91:254–66.
29. Gilworth G, Chamberlain A, Harvey A, Woodhouse A, Smith J, Smith G, Tennant A. Development of a Work Instability Scale for rheumatoid arthritis. Arthritis Rheum. 2003;49:349–54.
30. Tang K, Beaton DE, Lacaille D, Gignac MA, Zhang W, Anis AH, et al. The Work Instability Scale for Rheumatoid Arthritis (RA-WIS): does it work in osteoarthritis? Qual Life Res. 2010;19:1057–68.
31. Gilworth G, Emery P, Barkham N, Smyth NG, Helliwell P, Tennant A. Reducing work disability in Ankylosing Spondylitis – development of a work instability scale for AS. BMC Musculoskelet Disord. 2009;10:68.
32. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics. 1993;4:353–65.
33. Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, de Wit M, Dijkmans BA, Dougados M, Englbrecht M, Gogus F, Heiberg T, Hernandez C, Kirwan JR, Mola EM, Cerinic MM, Otsa K, Schett G, Scholte-Voshaar M, Sokka T, von Krause G, Wells GA, Kvien TK. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Ann Rheum Dis. 2011;70:935–42.
34. Kirwan JR, Reeback JS. Stanford **Health Assessment Questionnaire** modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol. 1986;25:26–9.
35. Wolfe F. 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. J Rheumatol. 2001;28:982–9.
36. Garrett S, Jenkinson T, Kennedy LJ, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in Ankylosing Spondylitis: The Bath Ankylosing Spondylitis Disease Activity Index (BASDAI). Journal of Rheumatology 1994:21;2286–2291.
37. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallory P, Jenkinson T. A new approach to defining functional ability in ankylosing spondylitis: The development of the Bath Ankylosing Spondylitis Functional Index (BASFI), Journal of Rheumatology 1994:21;2281–2285.
38. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. 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. J Rheumatol. 1988;15:1833–40.

39. Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. Arthritis Res Therapy. 2009;11(5):415.
40. Teresi JA, Kleinman M, Ocepek-Welikson K. Modern psychometric methods for detection of differential item functioning: application to cognitive assessment measures. Stat Med. 2000;19(11–12):1651–83.
41. Nunnally JC. Psychometric theory. New York: McGraw-Hill; 1978.
42. Rumm Lab. RUMM2030. https://www.rummlab.com.au. Accessed 6 June 2022.
43. IBM Corp. IBM SPSS Statistics for Windows, Version 26.0. Armonk. NY: IBM Corp; Released 2019.
44. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: The University of Chicago Press; 1980.
45. Tennant A, Conaghan PG. 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 Rheum. 2007;57(8):1358–62.
46. Evans JD. Straightforward statistics for the behavioural sciences. Pacific Grove (CA): Brooks/Cole Publishing; 1996.
47. Cichetti DV. Guidelines, criteria and rules of thumb for evaluating normed and standardised assessment instruments in psychology. Psychol Assess. 1994;6:284–90.
48. Donoghue D, group Stokes PROP. E. How much change is true change? The Minimum Detectable Change of the Berg Balance Scale in Elderly People. J Rehabil Med. 2009;41:343–6.
49. Stratford PW. Getting more from the literature: estimating the standard error of measurement from reliability studies. Physiother Can. 2004;56:27–30.
50. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were prosed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppMaterialsWALSpsychtestingJOR2022FINAL106.6.22.pdf