Received: 7 September 2023 | Accepted: 16 September 2023

DOI: 10.1002/msc.1829

WILEY

RESEARCH ARTICLE

Psychometric testing of the Rheumatoid Arthritis Work Instability Scale in employed people with fibromyalgia

Alison Hammond[1] | Alan Tennant[2] | Tamara Brown[1] | Yeliz Prior[1] | Angela Ching[1,3] | Jennifer Parker[1]

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

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

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

Correspondence

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

Funding information

University of Salford Manchester; European Alliance of Associations for Rheumatology, Grant/Award Number: HPR035; University of Salford

Abstract

Objective: The aim of the study was to establish whether the Rheumatoid Arthritis Work Instability Scale (RA-WIS), in its current form, is applicable for use with employed people with fibromyalgia (FM) to identify the risk of work disability and need for work rehabilitation.

Methods: Content validity was first investigated using cognitive debriefing interviews. Participants completed a postal questionnaire. Construct validity was assessed using Rasch analysis. Concurrent validity included testing between the RA-WIS and work (e.g., Workplace Activity Limitations Scale) and health (FM Impact Questionnaire-Revised (FIQ-R) scales. Two weeks later, participants were mailed a second questionnaire to measure test-retest reliability.

Results: Interviews were conducted with 13 participants with FM. All RA-WIS items were considered very or extremely relevant by almost all participants, with only one suggesting other items (anxiety and brain fog). Questionnaire responses were analysed from 156 employed participants: 94% women; 45.71 (SD 10.05) years of age; with time since FM diagnosis 2.99 (4.17) years (symptom duration 8.36 (SD 7.16) years). The RA-WIS mostly satisfied Rasch model requirements and a Rasch transformation scale was created. Concurrent validity was generally good (r_s = 0.55–0.66) with work scales and the FIQ-R. Internal consistency (Person Separation Index values) was consistent with group use in FM, not individual level use. Test-retest reliability was excellent, with intraclass coefficient (2, 1) = 0.90.

Discussion: The RA-WIS is valid and reliable for group use in employed people with FM. However, further work is needed to develop a WIS for individual use in FM.

KEYWORDS

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

1 | INTRODUCTION

Fibromyalgia (FM) is a common long-term condition affecting up 6.6% of the population worldwide and around 5% in the United Kingdom (UK). It affects more women than men, with peak onset during working age (25–55 years) (Fayaz et al., 2016; Marques et al., 2017). FM leads to chronic widespread pain, fatigue, muscle stiffness, non-restorative sleep, emotional distress, and cognitive issues, or brain fog, which affects memory, concentration, and thought processes (Salaffi et al., 2022). In the first year following diagnosis, work disability (i.e., stopping work prematurely due to ill-health) occurs rapidly in women with FM (from 60% employed to 41%), with many already having reduced or stopped work before diagnosis (Palstam & Mannerkorpi, 2017). For those working with FM, absenteeism has been estimated as 7%, presenteeism (i.e., at-work productivity loss) 44%, and overall work productivity loss (i.e., absenteeism and presenteeism) at 58% (Salaffi et al., 2022).

Given the impact of FM on employment, work rehabilitation and work modifications could allow more people with FM to remain employed, as well as alleviate symptoms experienced whilst working (Rakovski et al., 2012). One outcome measure which could be used to rapidly identify those benefiting from work interventions is the Work Instability Scale, developed for use in rheumatoid arthritis (RA) (Rheumatoid Arthritis Work Instability Scale (RA-WIS)) (Gilworth et al., 2003). Work instability is defined as a mismatch between an individual's functional abilities and the demands of their job, the consequences of which could threaten continuing employment if not addressed (Gilworth et al., 2003). The RA-WIS is a self-report measure with demonstrated reliability and validity in RA, including criterion validity with a full vocational assessment. With 23 dichotomous (true/false) items, it takes around five minutes to complete and be scored. Scores identify three levels of work instability indicating job loss risk: low (score 0–9), medium (10–17), and high (18–23) and the level of work rehabilitation needed, from simple to complex (Gilworth et al., 2003). The RA-WIS has also been shown to be suitable for use with working people with osteoarthritis, having good internal consistency, construct validity and responsiveness (Tang et al., 2010).

Before using the RA-WIS in FM, it is important to identify whether working people with FM consider RA-WIS items reflect their experiences in relation to work (content validity) and establish its psychometric properties in FM. Testing should include both classical testing and item response theory (e.g., Rasch analysis) to establish reliability and validity. The aim of the study was to establish whether the RA-WIS, in its current form, is applicable for use with working people with FM.

2 | METHODS

2.1 | Design, participants, and recruitment procedures

The study consisted of two phases. In Phase 1, cognitive debriefing interviews were conducted to investigate content validity and ease of completion of the RA-WIS from employed people with FM's perspective (De Vet et al., 2011). In Phase 2, cross-sectional surveys were conducted to establish the psychometric properties of the RA-WIS in FM. 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. Ethical approval was obtained from the National Research Ethics Service Committee East Midlands—Leicester South (17/EM/0409). All participants provided written, informed consent.

Participants were recruited using convenience sampling: in Phase 1 from four secondary care National Health Service (NHS) Trusts' rheumatology out-patient clinics; and in Phase 2, from 30 secondary care and four community NHS Trusts' Rheumatology or Therapy out-patient clinics, with some from a University Arthritis Volunteer Register, in the UK. Participants were eligible if at least 18 years old; in paid employment at least 1 day a week; currently working, or on less than 4 weeks of sick leave with entry delayed until returned to work; and a primary diagnosis of FM (i.e., not secondary to other forms of arthritis, such as RA). Diagnosis was confirmed by a rheumatologist, general practitioner, or extended scope physiotherapist. Participants needed to be able to read, write and understand English. Patients were ineligible if on long-term sick leave because the RA-WIS asks about work "at the moment." Patients were identified by therapists or research facilitators using these criteria and given a short study explanation and information pack. The latter included a participant information sheet, Freepost envelope and reply form (including diagnosis, and employment and sick leave status, to check eligibility criteria) to return to the research team. This study was conducted alongside testing of several work outcome measures (the WORK-PROM study). Full recruitment procedures are detailed elsewhere (Hammond et al., 2023b).

2.2 | Data collection

In Phase 1, participants were mailed a paper questionnaire booklet, including the RA-WIS, to complete at home, and asked to consider the ease of completion, item relevance and if any important impacts of FM on work ability were missing. Within 2 weeks, they were interviewed, by telephone, about comprehensiveness (1 = not relevant; 5 = extremely relevant; and any missing items) and comprehensibility (instructions, content, layout).

In Phase 2, following receipt of the reply form, the research team mailed a paper questionnaire booklet to the patient to complete at home (Test 1: T1). The front page of this included a consent form to complete. Two weeks after T1 return, they were mailed a second questionnaire (Test 2: T2) to assess test-retest reliability. Following each mailing, if required, participants were sent a reminder at 2 weeks (letter) and 4 weeks (letter plus questionnaire booklet).

The T1 booklet included demographic data: 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).

2.2.1 | Work and health scales

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

As well as the RA-WIS, three scales evaluated both the physical and psychological impact of health conditions on work. These were the British-English version of the Workplace Activity Limitations Scale (WALS), a measure of presenteeism, with 12 items of physical work ability (eight items); managing work demands (physical and/or mental) (three items); and concentration at work (one item), scored 0 = no difficulty to 3 = unable to do (range 0–36) (Hammond et al., 2023a). The Work Limitations Questionnaire-25 indicates the percentage of productivity loss in the past two weeks (Lerner et al., 2001). The Work Productivity and Activity Impairment Work Productivity Activity Impairment (General Health) scale, which includes six items from which a Percentage Overall Work Impairment due to Health (in the past 7 days) score is calculated (Reilly et al., 1993). Additionally, work self-efficacy, measured on a 0–10 numeric rating scale, was collected, with higher scores indicating greater self-efficacy.

The Revised Fibromyalgia Impact Questionnaire was also included: overall impact (two items: score range 0–20), symptoms (10 items: 0–50); and physical function (nine items: 0–30). This provides a FIQ-R score with a range 0–100 (Bennett et al., 2009). 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 RA-WIS again plus perceived change in health status 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

In Phase 1, at least 10 participants should be included in cognitive debriefing interviews (Terwee et al., 2012). In Phase 2, a sample size between 100 and 500 is sufficient for testing scales with dichotomous data (e.g., true/false) if well targeted, as Rasch analysis was used to assess construct (structural) validity (Hagell & Westergren, 2016; Linacre, 1994; Rasch, 1980). At least 79 sets of repeated responses were needed to demonstrate that a test-retest correlation of 0.70 differs from a background correlation (constant) of 0.45, with 90% power at the 1% significance level. A test-retest reliability correlation of 0.70 is considered a minimum acceptable level (Nunnally, 1978).

2.4 | Statistical analyses

Demographic, 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 using the dichotomous model (explained in construct validity, below) (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 v29 (IBM Corp, 2023). 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. Less than 3% of missing data is acceptable and more than 15% unacceptable (De Vet et al., 2011).

2.4.2 | Validity

Construct (structural) validity and Rasch Measurement Theory. In 1960, Georg Rasch first published his book "Probabilistic Models for some Intelligence and Attainment Tests" (Rasch, 1960). Here, he introduced a model for dichotomous responses in a test, subsequently written as an "additive logistic model," which is the form widely reported today (Kreiner, 2013). The model states that the probability of obtaining a correct response to an item, given the ability level of the person, is a logistic function of the difference between the ability of the person and the difficulty of the test item (Tennant & Küçükdeveci, 2023). In the context of work instability, this would mean that the probability of a person with FM affirming an item on the RA-WIS would represent the difference between their level of work instability and the level of instability represented by the item.

A further attribute of the Rasch model is that it can uncover the levels of instability of the items in a scale, providing a hierarchical ordering of items from low (instability) impact to high impact. Within this model, it is expected that only a person with high levels of work instability would be likely to affirm an item with high impact on their work. Rasch analysis sets out to test if this is the case. This aspect is usually tested by some form of Chi-Square test, indicating no difference between the response to the item, and that expected by the model. A 5% significance level is applied in this study, Bonferroni adjusted (i.e., to reduce the likelihood of results incorrectly appearing statistically significant when multiple tests are conducted). This forms one part of a series of tests to see if the data from the scale accord with the Rasch model's expectations.

Further testing incudes local item independence (Marais & Andrich, 2008). This is where items should be uncorrelated after conditioning on the total score. Recent work has shown that the threshold for identifying items that are locally dependent is where the residual correlations of items are ≥0.20 *above the average* residual

correlation (Christensen et al., 2017). This is applied in this study. If local item dependency is found, items are merged (super items) and the fit to the model retested (Baghaei, 2010).

Testing also includes group invariance. That is, for those at the same level of work instability, the response to an item should be the same, irrespective of group membership, for example, their age or level of education. This is referred to as Differential Item Functioning (DIF) (Hagquist & Andrich, 2017; Tennant et al., 2004; Teresi et al., 2000). It was tested in this study through an analysis of variance of residuals. Group membership was tested across several contextual factors, including age, time since symptom onset and since diagnosis, job skill level and educational level. Once again, a 5% significance level is applied, Bonferroni adjusted. If an item is found to have a significant difference, consideration must be given to splitting the item in some way (e.g., by education level) to see if the problem can be resolved.

All summative scales, such as the RA-WIS, should be unidimensional. In this study, this is tested by Smith's test approach, where items are differentiated on the residuals into positive and negative sets, and the two sets of estimates are compared to see if they are different (Smith, 2002). Once again, a 5% significance level is applied. A significant difference indicates some level of multidimensionality.

Concurrent validity: This is the degree to which scale scores correlate with other relevant scales. This was assessed using Spearman's correlations. Correlations of 0.20–0.39 are considered weak, 0.4–0.59 moderate, and ≥0.6 strong (Evans, 1996). Moderate to strong positive correlations were hypothesised between the RA-WIS, work and health scales.

Discriminant validity: This refers to hypothesis testing that there will be significant score differences in RA-WIS scores between those reporting good, fair, and poor health. This was assessed using Kruskal-Wallis tests, with $p \leq 0.05$ considered significant.

2.4.3 | Reliability

Internal consistency: This is the degree of interrelatedness between items within a scale. This can be assessed using (a) Cronbach's alpha. For which results ≥0.80 are deemed good to excellent, with ≥0.90 being consistent with individual use; and >0.70 with group-level use (Evans, 1996); and (b) the Person Separation Index (PSI) for which scores >0.70 also indicate group-level use; and ≥0.85 individual use (Tennant & Küçükdeveci, 2023).

Test-retest reliability: This refers to the extent to which scores are the same for repeated measurements over time in those reporting that their health has not changed (i.e., for whom perceived health is "the same" at T1 and 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 0.50–0.74 moderate (Cichetti, 1994). Reliability of individual scale items was calculated using 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 formula (SEM x 1.96 x $\sqrt{2}$). It is a statistical estimate of the SDD across groups above measurement error (Donoghue, PROP group & Stokes, 2009).

Floor and ceiling effects: These 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

Cognitive debriefing interviews were conducted with 13 participants with FM. Demographic and work characteristics are in Table 1. All items were considered very or extremely relevant by almost all participants, with only one to three participants considering 12 items as not, a little or fairly relevant, with most such responses from two participants (Supplementary Table 1). All 13 participants considered that the main work issues relevant to them were included in the RA-WIS. Only one participant suggested two additional items: "*anxiety is a big issue in fibromyalgia, either due to the job, the effects on physical health or something outside of work*" and "*brain fog, as it often means having to do certain things at certain times, such as having to leave a task to another day*" (woman, 57years, FM 15years, civil servant). The first suggested item, anxiety, is partly addressed by item 9 (worry about ability to keep working). However, there are no items directly addressing cognitive issues (brain fog), although this could affect responses to items 3, 6, 16, 18, and 19. As only one participant suggested extra items, and the aim of the study was to identify if the RA-WIS is applicable in FM, additional items were not added. All stated the RA-WIS was easy to understand and complete.

3.2 | Phase 2

Overall, 264 people with FM were referred to the study of whom 161 completed questionnaires (i.e., a 61% response rate): secondary care n = 135/209 (64.60% response rate), community care 23/51 (45.10%) and 3 of the 4 volunteers (75%). Five were excluded: four reported in their questionnaire that RA or axial spondyloarthritis was their primary diagnosis, and one because unemployed. Accordingly, 156 T1 questionnaires were analysed, and 115 participants also returned T2 questionnaire booklets (73.72%). Participant characteristics are shown in Table 1 and their work and health measures in Table 2.

TABLE 1 Demographic and work characteristics of participants: Phase 1 and 2.

	FM	
	Phase 1	**Phase 2**
n =	13	156
Sex M:F, *n* (%)	2:11	10 (6): 146 (94)
Age (years), mean (SD)	39.69 (9.11)	45.71 (10.05)
Job skill level: *n* (%)		
1 and 2	7	95 (61)
3 and 4	6	61 (39)
Disease duration (years), mean (SD)	5.38 (3.55)	2.99 (4.17)
Phase 2 only		
Symptom duration (years), mean (SD)	8.36 (7.16)	
Living with spouse/family/significant other, *n* (%)	139 (89)	
Children <18years living at home, *n* (%)	56 (36)	
Educational level (ISCED), *n* (%):		
No formal qualifications	7 (4)	
Secondary/non tertiary	76 (49)	
Tertiary	73 (47)	
Full- or part-time work, *n* (%)	70 (45):86 (55)	
Hours worked, mean (SD)	31.50 (10.56)	
Self-employed, *n* (%)	18 (11)	
Physical demands of job, *n* (%):		
None/a little	61 (39)	
Noticeable	14 (9)	
A lot/great deal	81 (52)	
Medication regimen, *n* (%):		
None	23 (15)	
NSAIDS +/− analgesics	14 (9)	
Steroids +/- NSAIDS	6 (4)	
Neuropathic analgesics (e.g., gabapentin/pregabalin)	99 (64)	
Opiate medication	12 (8)	

Abbreviations: FM, fibromyalgia; ISCED, International Standard Classification of Education; NSAID, non-steroidal anti-inflammatory drugs; ONS, Office for National Statistics; SD, standard deviation.

3.2.1 | Compliance

Missing data were very low, with only two items unanswered by one participant, who wrote these were "not applicable" (<0.001%), and items therefore re-scored as 0.

3.2.2 | Validity

Construct (structural) validity: Data for the 23 dichotomous items were fit to the Rasch model. Initial fit was poor, showing a significant Chi-Square test of fit, although all other indicators were satisfactory, including the absence of DIF on the contextual factors (Table 3 see Baseline). The item most easily affirmed was item 6, "I get good days and bad days at work", while the item least likely to be affirmed was item 17, "I have great difficult opening some of the doors at work."

One item in particular, "I push myself to go to work because I don't want to give in to my condition" (item 13) appeared to misfit (Table 4). However, this also displayed local dependency with item 7, "I can get my job done, I'm just a lot slower", and so the two items were merged to form a super-item. In all, three clusters of locally dependent items were observed, involving eight of the 23 items. All three of the clusters were grouped into super items, and the data refitted to the model. This improved fit gave a satisfactory result,

TABLE 2 Phase 2: Participants' work and health measures.

	FM (*n* = 156)
Work measures	
RA-WIS (0–23), median IQR):	18.00 (15.00–20.00)
Low work instability (0–9), *n* (%)	6 (3.84)
Moderate work instability (10–17), *n* (%)	64 (41.03)
High work instability (18–23), *n* (%)	86 (55.16)
WALS (0–36), median (IQR)	16.00 (12.00–19.00)
WLQ-25 (0–100), median (IQR):	
Time management demands	60.00 (40.00–80.00)
Physical demands	58.33 (43.75–73.96)
Mental interpersonal demands	44.44 (27.78–61.11)
Output demands	45.00 (25.00–65.00)
WLQ-25 % productivity loss	13.26 (9.20–16.53)
WLQ-25 summed score	51.69 (37.30–64.62)
WPAI, median (IQR):	
% Overall work impairment due to health	66.15 (50.00–80.00)
Work self-efficacy (0–10), median (IQR),	7.00 (5.00–8.00)
Health measures	
Perceived severity health last month (1–5; median IQR) *n* (%)	4.00 (3.00–4.00)
Poor/very poor	83 (53)
Fair	63 (41)
Good/very good	10 (6)
FIQ-R (normalised scores: median IQR):	
Overall impact (0–20)	14.00 (10.00–17.00)
Symptoms (0–50)	34.50 (28.13–39.00)
Function (0–30)	19.33 (14.67–22.67)
FIQR total (0–100)	68.33 (54.20–77.50)
Completed T1 and T2 questionnaires:	*n* = 116
Time between T1 and T2, days (median, IQR)	35 (27–47)
Self-reported effect of health condition at T2 versus T1, *n* (%):	
Much/somewhat less troublesome	14 (12)
The same	54 (47)
Somewhat/much more troublesome	48 (41)

Note: For all measures: higher scores indicate more work/health problems (except for work self-efficacy, for which higher scores indicate greater self-efficacy).

Abbreviations: FIQ-R, Fibromyalgia Impact Questionnaire—Revised; FM, fibromyalgia; IQR, inter-quartile range; NRS, numeric rating scale; RA-WIS, Rheumatoid Arthritis-Work Instability Scale; SD, standard deviation; WALS, Workplace Activity Limitations Scale; WLQ-25, Work Limitations Questionnaire-25; WPAI, Work Productivity Activity Impairment.

including unidimensionality, the absence of any DIF, and retaining 95% of the variance after absorbing the effect of local item dependency (Table 3—see Adjusted for local item dependency).

Irrespective of the fit, the targeting of the scale was less than ideal (Figure 1). Participants with FM had much higher levels of work instability (Logit 1.40) than the average of the scale itself (Logit Zero). This may have influenced the reliability (internal consistency), as Cronbach's **α** overestimates reliability in mistargeted data, and so the PSI is a more relevant statistic. There was no DIF, that is, no significant differences in the levels of work instability by groups of age,

TABLE 3 Summary of fit of the RA-WIS in fibromyalgia to the Rasch model.

	Residuals		Chi-Square		Reliability		Dimension	DIF	ECV
Analysis	**Item**	**Person**	**Value (df)**	***p***	**PSI**	**α**	**% *t*-tests (LCI)**		
Baseline	0.86	0.77	74.50 (46)	0.005	0.73	0.80	3.90	None	-
Adjusted for local item dependency	0.82	0.74	46.00 (36)	0.09	0.69	0.76	3.20	None	0.95
Ideal values	**1.0**	**1.0**		**>0.05**	**>0.70**	**>0.7**	**<5%**	**None**	**>0.9**

Note: Ideal values are those one expects if data fit the Rasch model.

Abbreviations: DIF, Differential Item Functioning; ECV, Explained Common Variance; LCI, Lower Confidence Interval; PSI, Person Separation Index; SD, Standard Deviation; α, Cronbach's alpha.

TABLE 4 Individual item fit of the rheumatoid arthritis (RA) work instability scale in fibromyalgia.

Item		Location (Logits)	Residual	Chi-Square	DF	P
1	I'm getting up earlier because of my condition	0.85	1.63	7.37	2	0.03
2	I get very stiff at work	−1.96	−0.72	0.67	2	0.72
3	I'm finding my job is about all I can manage	−0.43	−0.12	0.42	2	0.81
4	The stress of my job makes my condition flare	0.19	−0.65	1.50	2	0.47
5	I'm finding any pressure on my hands is a problem	0.86	0.47	0.70	2	0.71
6	I get good days and bad days at work	−2.26	0.46	7.79	2	0.09
7	I can get my job done, I'm just a lot slower	0.36	0.22	1.21	2	0.55
8	If I don't reduce my hours, I may have to give up work	2.15	−1.41	4.69	2	0.10
9	I am very worried about my ability to keep working	−0.42	−2.03	9.76	2	0.01
10	I have pain or stiffness all the time at work	−0.27	−0.31	0.59	2	0.74
11	I don't have the stamina to work, like I used to	−0.66	0.69	8.03	2	0.02
12	I have used my holiday so that I don't have to go sick	1.87	0.35	1.15	2	0.56
13	I push myself to go to work because I don't want to give in to my condition	−1.95	0.70	13.89	2	0.001
14	Sometimes I can't face being in work all day	−0.99	0.19	0.38	2	0.83
15	I have to say no to certain things at work	0.49	0.30	3.27	2	0.20
16	I've got to watch how much I do certain things at work	−0.16	−0.67	0.82	2	0.67
17	I have great difficulty opening some of the doors at work	2.38	0.21	0.29	2	0.87
18	I have to allow myself extra time to do some jobs	−0.40	−0.39	1.04	2	0.60
19	It's frustrating because I can't always do things at work	0.10	−1.78	4.91	2	0.09
20	I feel I may have to give up work	1.34	−0.60	2.43	2	0.30
21	I get on with the work but afterwards I have a lot of pain	−1.56	−1.22	2.23	2	0.33
22	When I'm feeling tired all the time work's a grind	−1.12	−0.69	0.96	2	0.62
23	I'd like another job, but I am restricted to what I can do.	1.58	−0.53	3.42	2	0.18
Ideal values			**Within ±2.5**			**0.05 Bonferroni Adjusted**

Note: Bonferroni Adjustment $p = 0.002$. Ideal values are those one expects if data fit the Rasch model.

duration, education, sex, job skill level or whether self-employed. A transformation table for converting RA-WIS raw scores to a metric scale (i.e., to convert data to be suitable for use in parametric tests) is given in Table 5.

Concurrent validity: As hypothesised, the RA-WIS exhibited moderate to strong correlations with work measures (r_s = 0.55–0.66) (except for the WLQ Physical Demands sub-scale, which was weak (r_s = 0.32), and with the FIQ-R (r_s = 0.56) (Table 6).

Discriminant validity: There were significant differences in RA-WIS scores between those with differing levels of self-reported health: poor/very poor 19.00 (17.00–20.00) (n = 83); fair 17.00

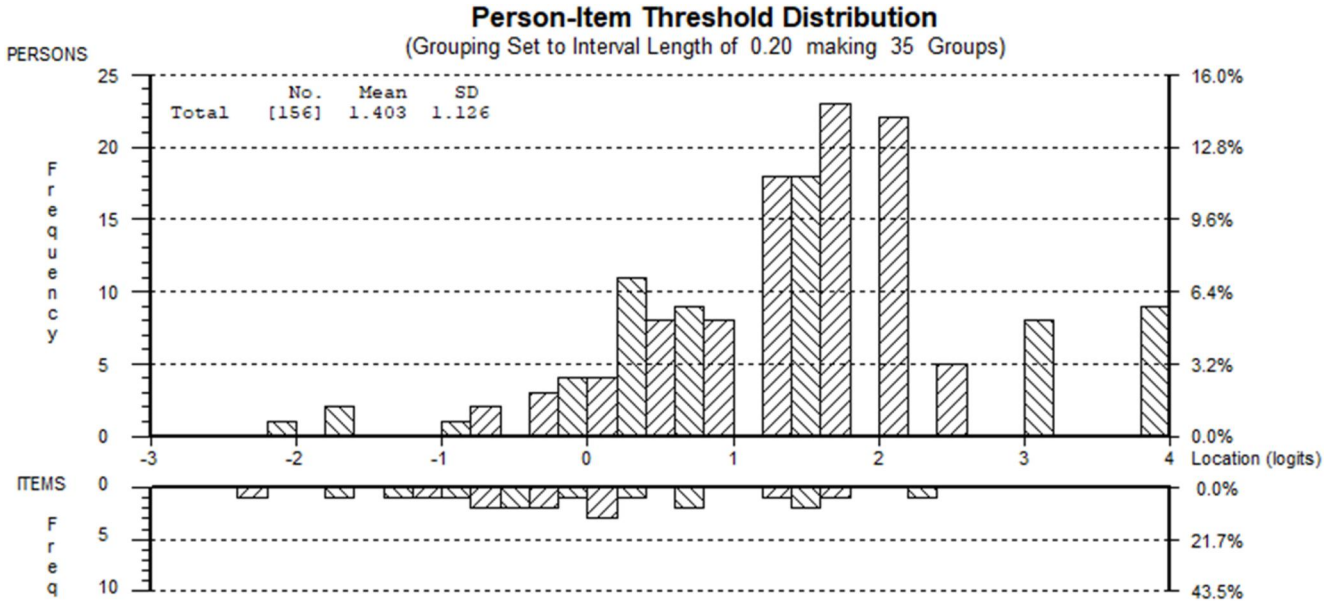

FIGURE 1 Item-Person Distribution of rheumatoid arthritis (RA) Work Instability Scale in fibromyalgia.

TABLE 5 Transformation table for the RA-WIS in fibromyalgia.

Raw score	Metric
0	0.0
1	2.6
2	4.4
3	5.7
4	6.7
5	7.5
6	8.2
7	8.8
8	9.5
9	10.0
10	10.6
11	11.2
12	11.8
13	12.4
14	13.0
15	13.7
16	14.3
17	15.0
18	15.8
19	16.6
20	17.5
21	18.7
22	20.5
23	23.0

(13.00–19.00) (n = 63); and good/very good 13.50 (9.25–16.50) (n = 10): H = 27.17; df = 2; $p < 0.001$.

3.2.3 | Reliability

Internal consistency: Cronbach's' alpha = 0.78 and PSI = 0.76 (−adjusted).

Test-retest reliability: At T2, 54 participants reported their health was "the same" as at T1 and were included in the analysis. Median scores were the same at T1 and T2, at 18.00 (15.00–20.00), with a strong correlation between T1 and T2 of $r_s = 0.67$ ($p < 0.001$). The ICC (2,1) = 0.90 (95% CI: 0.83, 0.94), that is, excellent. Item reliability was mostly moderate to good, with six items weak (items 2, 6, 11, 13, 14 and 21) (supplementary table 2).

3.2.4 | Precision

Precision. The SEM = 1.20, and SDD = 3.38.

Floor and ceiling effects. None scored 0, and 9/156 (5.77%) scored 23, that is, within acceptable limits (<15%).

4 | DISCUSSION

This is the first study to investigate whether the RA-WIS is applicable for use in working people with FM. Content validity was demonstrated from the perspective of working people with FM. The RA-WIS was found to satisfy Rasch model expectations in FM after accommodating some local item dependency. Nevertheless, the scale had poor targeting and internal consistency was less than that required for individual use (PSI ≥0.85). As such, the RA-WIS is only

TABLE 6 Concurrent validity of the rheumatoid arthritis (RA) Work Instability Scale with work and health measures in fibromyalgia.

RA-WIS correlations with:	FM (n = 156) r_s
Work measures	
WLQ-25 (0–100):	
Time management demands	0.61**
Physical demands	0.32**
Mental interpersonal demands	0.61**
Output demands	0.58**
WLQ-25% productivity loss	0.66**
WLQ-25 summed score	0.66**
WALS	0.60**
WPAI (%)	
Overall work impairment due to health	0.56**
Work self-efficacy	0.55**
Health measures	
FIQR (normalised scores):	
Overall impact (0–20)	0.47**
Symptoms (0–50)	0.54**
Function (0–30)	0.40**
FIQ-R total (0–100)	0.56**

Note: **, correlation significant at 0.01 level.

Abbreviations: FM, fibromyalgia; FIQ-R, Fibromyalgia Impact Questionnaire—Revised; RA-WIS, Rheumatoid Arthritis Work Instability Scale; rs, Spearman's correlation; WALS, Workplace Activity Limitations Scale; WLQ-25, Work Limitations Questionnaire-25; WPAI, Work Productivity Activity Impairment.

suitable for group level use, for example, research in FM, whereas it was designed for clinical and research use in RA (Gilworth et al., 2003). Given the level of reliability, additional work on scale development is needed to raise the level to be consistent with clinical use (Bland & Altman, 1997). Confirmation is also needed that the original RA-based cut points for work instability remain valid.

These findings suggest the need for further exploration of potential items reflecting more severe levels of work instability in working people with FM. The RA-WIS predominantly focuses on items related to symptoms (e.g., pain and fatigue), physical and mental job demands, self-efficacy and job abilities. Phase 1 participants endorsed these items as very relevant. A systematic review confirms their relevance, as common problems reported by working people with FM are pain and fatigue causing work difficulties, reduced work self-efficacy and job competency (Mukhida et al., 2020). However, research also identifies other factors not captured in the RA-WIS impacting work instability in FM. These include the negative impact of FM on occupational identity (Chen et al., 2019; Depelteau et al., 2021; Mukhida et al., 2020); motivation at work (Arnold et al., 2008); reduced social support from friends, and strained relationships at work (Depelteau et al., 2021; Mukhida et al., 2020); as well as concerns about disclosing their condition at work and social stigma associated with FM from colleagues and supervisors (Mukhida et al., 2020). A recent study identified that FM-related sick leave was not directly related to the condition itself, but rather issues such as commute time, repetitive work movements, work difficulties, job stress, noisy workplaces, limited career progression, and lack of recognition or understanding of FM by managers and co-workers (Laroche et al., 2019). The comments from one Phase 1 participant suggested that RA-WIS should also include anxiety and brain fog. The above suggests that, for clinical use, an FM-WIS is needed which captures the most important range of symptoms, work difficulties, personal, and work environment (in the widest sense) factors and better reflects people with FM's working experiences.

Using the RA-derived clinical cut points, in this sample of working people with FM recruited from hospital and community outpatient clinics, only 3.90% had low work instability, 41.60% had moderate and 54.50% high work instability, that is, almost all were at risk of work disability and in need of work rehabilitation. The average time between symptom onset and diagnosis was 5 years, which was similar to that found elsewhere (Collado et al., 2014). Within 5 years of FM symptom development, 24% stop working and 33% reduce paid work hours (Guymer et al., 2016). Participants in this sample therefore represent those who have managed to continue working. Over half worked part-time (55%), compared to 38% of working women in the general UK population (Buchanan et al., 2023), suggesting that many had already reduced working hours. Between 70% and 94% of participants in this sample reported difficulties performing the 12 activities included in the WALS but only 9%–31% had work adaptations for these, indicating many were not getting work support (Brown et al., 2023).

To help people with FM keep working, it is essential that the time between symptom onset and diagnosis is reduced, work-related needs are identified quickly, and work support services provided earlier. The UK's "Getting It Right First Time" programme recommends that diagnosis and management of people with FM are based in primary care (Wilson et al., 2022). Diagnosis of FM can be difficult, because of the wide-ranging symptoms overlapping with other conditions, and people with FM report limited support from health services (Wilson et al., 2022). Diagnostic delay is likely influenced by low levels of skill and confidence in diagnosing FM among health professionals, especially general practitioners (GPs), leading to referrals for diagnosis (e.g., to rheumatology or pain consultants), as well as reluctance by GPs to label people with this diagnosis (Wilson et al., 2022). Recent diagnostic guidelines for FM from the Royal College of Physicians have been designed to address this diagnostic delay and emphasise providing information and rehabilitation in parallel to investigations (Berwick et al., 2022).

The National Institute for Health and Care Excellence (NICE) guidelines for the management of chronic primary pain (such as FM) emphasise identifying how chronic pain is affecting people's lives, including work (NICE, 2021). Exercise and physical activity,

psychological therapies, and pharmacological management of chronic pain are recommended, but work rehabilitation is not mentioned. Screening for work problems should be introduced while the condition is under investigation, people with FM should be provided with information about managing work issues, and work rehabilitation, if needed. An FM-WIS could assist in identifying physical and mental work difficulties, as well as personal and work environment factors affecting work. GPs can issue Fit Notes, for people on sick leave or working but experiencing difficulties, recommending to employers which work modifications can help people keep working. However, of over 6 million Fit Notes issued by GPs in 2020 for adults on sick leave (for any cause), 95% did not suggest work adjustments or advice (Shemtob & Asanati, 2021). Musculoskeletal first contact practitioners located in GP practices, and health professionals working in community and secondary care (occupational therapists, physiotherapists, nurses, pharmacists) can also complete Fit Notes. These health professionals are well-placed to assess for and recommend work adaptations, and cross-refer to other members of the multidisciplinary team for work rehabilitation, as appropriate (Department for Work and Pensions, Department of Health and Social Care, 2022; NHS England & NHS Improvement, 2019; Nouri et al., 2021). This could help reduce the high costs to individuals and to society from the absenteeism, presenteeism and work disability associated with FM (D'Onghia et al., 2022).

4.1 | Limitations, and future research

Participants were recruited from a wide variety of NHS out-patient clinics, meaning that the results are representative for people accessing secondary or community care. Half were diagnosed with FM for a year or less, but most had symptoms for some years, and already had medium to high levels of work instability. The targeting of the scale identified participants had higher levels of work instability than the average of the scale. More participants with recent onset symptoms would have increased representativeness. Recruiting from primary care, which may have helped with this, was not feasible as the Clinical Research Networks supporting the study were unable to identify any to assist (National Institute of Health Research, 2023). Additionally, very few men with FM were recruited, although most diagnosed with FM (80%–96%) are women (Ruschak et al., 2023). A further limitation was the sample size, due to the poor targeting of the scale. It is possible that disturbances to model fit remain undiscovered due to this, meaning the sample needed to be larger. Also, the smaller sample size at T2 precluded cross validation between T1 and T2 results, which may have strengthened the results. This was also a smaller sample for test-retest reliability than required, as relatively few reported stable health at T2.

Future research could focus on developing a WIS for FM, incorporating a wider range of personal and work environment factors influencing work participation in FM, based on literature review and in-depth interviews. In Phase 1, whilst 13 participants were deemed sufficient for cognitive debriefing interviewing, a larger number could provide better insights when developing and testing items for inclusion. The study identified the need for items reflecting more severe levels of work instability. This sample should therefore include participants with high levels of work instability, and those already having stopped work due to FM to ensure such items are generated. An FM-WIS would also need criterion testing against workplace assessments.

4.2 | Conclusion

Overall, psychometric testing of the RA-WIS in FM demonstrated acceptable validity and reliability in employed people with FM, but for group-level use only. Accordingly, the RA-WIS can be used in the UK in FM for research and other group-level studies. The study met most recommendations of the COSMIN checklist for methodological quality and reporting (Gagnier et al., 2021; Mokkink et al., 2010). Transformation to a metric scale is available for calculations of change and other parametric procedures, distributions permitting.

AUTHOR CONTRIBUTION

Alison Hammond and Alan Tennant designed the study. Phase 1: Alison Hammond and Yeliz Prior conducted interviews. Phase 2: Material preparation and data collection were performed by Angela Ching, Jennifer Parker, and Alison Hammond. Analyses were performed by Alan Tennant (Rasch analysis) and Alison Hammond (Phase 1 and classical testing). The first draft of the manuscript was prepared by Alison Hammond, Alan Tennant and Tamara Brown. All authors commented on previous versions of the manuscript and read and approved the final manuscript.

ACKNOWLEDGEMENTS

Many thanks to: Dr Mike Horton, Leeds Psychometric Laboratory for Health Sciences, University of Leeds, UK, for permission to test the RA Work Instability Scale in fibromyalgia; to the many participants with fibromyalgia who completed the interviews and study questionnaires; and 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: Karen Crosby (LC), JoAnn Nicholson, Susannah Glasgow, Manchester University NHS Foundation Trust; Nicky Walker (LC), Mid Cheshire Hospitals 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; 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; 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 Imran Riaz (LC), Ellie Gilham, Jayne Brown, The Queen Elizabeth Hospital King's Lynn 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; 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; 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. NORTHERN IRELAND: Una McKenna (LC), Northern Health and Social Care Trust. This work was conducted alongside the WORK-PROM study (part-funded by the European Alliance of Associations for Rheumatology (EULAR) grant number HPR035), with additional funding from the University of Salford, and 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 Yeliz Prior (y.prior@salford.ac.uk). 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 School of Health & Society 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
Tamara Brown https://orcid.org/0000-0003-1069-4627
Yeliz Prior https://orcid.org/0000-0001-9831-6254
Angela Ching https://orcid.org/0000-0002-3765-2534
Jennifer Parker https://orcid.org/0000-0002-2235-5748

REFERENCES

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.

Arnold, L. M., Crofford, L. J., Mease, P. J., Burgess, S. M., Palmer, S. C., Abetz, L., & Martin, S. A. (2008). Patient perspectives on the impact of fibromyalgia. *Patient Education and Counselling*, *73*(1), 114–120. https://doi.org/10.1016/j.pec.2008.06.005

Baghaei, P. (2010). A comparison of three polychotomous Rasch models for super-item analysis. *Psychological Test and Assessment Modelling*, *52*, 313–323.

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

Berwick, R., Baker, C., & Goebel, A., & Guideline development group. (2022). The diagnosis of fibromyalgia syndrome. *Clinical Medicine*, *22*, 570–574. https://doi.org/10.7861/clinmed.2022-0402

Bland, J., & Altman, D. (1997). Statistics Notes: Cronbach's alpha. *British Medical Journal*, *314*(7080), 572. https://doi.org/10.1136/bmj.314.7080.572

Brown, T., Hammond, A., Ching, A., & Parker, J. (2023). Work limitations and associated factors in rheumatoid arthritis, axial spondyloarthritis, osteoarthritis and fibromyalgia. *Musculoskeletal Care*, *21*(3), 827–844. https://doi.org/10.1002/msc.1760

Buchanan, I., Pratt, A., & Francis-Devine, B. (2023). *Research briefing: Women and the UK economy*. UK Parliament, House of Commons Library. Retrieved from https://researchbriefings.files.parliament.uk/documents/SN06838/SN06838.pdf

Chen, A. T., Carpenter, H., & Flaherty, M. G. (2019). Work experiences, accommodations, and information in the context of fibromyalgia: A literature review and conceptual synthesis. *The International Journal of Information, Diversity, & Inclusion*, *3*(3). Special Issue Health Justice Part 1. https://doi.org/10.33137/ijidi.v3i3.32964

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, *41*(3), 178–194. https://doi.org/10.1177/0146621616677520

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

Collado, A., Gomez, E., Coscolla, R., Sunyol, R., Solé, E., Rivera, J., Altarriba, E., Carbonell, J., & Castells, X. (2014). Work, family, and social environment in patients with fibromyalgia in Spain: An epidemiological study: EPIFFAC study. *BMC Health Services Research*, *14*(1), 513. https://doi.org/10.1186/s12913-014-0513-5

Department for Work and Pensions, Department of Health and Social Care. (2022). Who can issue fit notes: Guidance for healthcare professionals and their employers. Retrieved from https://www.gov.uk/government/publications/who-can-issue-fit-notes-guidance-for-healthcare-professionals-and-their-employers/who-can-issue-fit-notes-guidance-for-healthcare-professionals-and-their-employers

Depelteau, A., Lagueux, E., Page, R., & Hudon, C. (2021). Occupational adaptation of people living with fibromyalgia: A systematic review

and thematic synthesis. *American Journal of Occupational Therapy*, *75*(4), 7504190040. https://doi.org/10.5014/ajot.2021.047134

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

D'Onghia, M., Ciaffi, J., Ruscitti, P., Cipriani, P., Giacomelli, R., Ablin, J. N., & Ursini, F. (2022). The economic burden of fibromyalgia: A systematic literature review. *Seminars in Arthritis and Rheumatism*, *56*, 152060. https://doi.org/10.1016/j.semarthrit.2022.152060

Donoghue, D., Stokes, E., & 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.

Fayaz, A., Croft, P., Langford, R. M., Donaldson, L. J., & Jones, G. T. (2016). Prevalence of chronic pain in the UK: A systematic review and meta-analysis of population studies. *BMJ Open*, *6*(6), e010364. https://doi.org/10.1136/bmjopen-2015-010364

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. Retrieved from https://www.cosmin.nl/wp-content/uploads/COSMIN-reporting-guideline_1.pdf

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 & Rheumatism*, *49*(3), 349–354. https://doi.org/10.1002/art.11114

Guymer, E. K., Littlejohn, G. O., Brand, C. K., & Kwiatek, R. A. (2016). Fibromyalgia onset has a high impact on work ability in Australians. *Internal Medicine Journal*, *46*(9), 1069–1074. https://doi.org/10.1111/imj.13135

Hagell, P., & Westergren, A. (2016). Sample size and statistical conclusions from tests of fit to the Rasch model according to the Rasch Unidimensional Measurement Model (RUMM) program in health outcome measurement. *Journal of Applied Measurement*, *17*(4), 416–431.

Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and Quality of Life Outcomes*, *15*(1), 181. https://doi.org/10.1186/s12955-017-0755-0

Hammond, A., Tennant, A., Ching, A., Parker, J., Prior, Y., Gignac, M. A. M., Verstappen, S., & O'Brien, R. (2023a). 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

Hammond, A., Tennant, A., Ching, A., Parker, J., Prior, Y., Gignac, M. A. M., Verstappen, S. M. M., & O'Brien, R. (2023b). Psychometric testing of the British-English perceived workplace support scale, work accommodations, benefits, policies and practices scale, and work transitions Index in four rheumatic and musculoskeletal conditions. *Musculoskeletal Care*. https://doi.org/10.1002/msc.1807

IBM Corp. (2023). *IBM SPSS statistics for windows*. Version 29.0. IBM Corp. Released.

Kreiner, S. (2013). The Rasch model for Dichotomous items. In K. B. Christensen, S. Kreiner, & S. Mesbah (Eds.), *Rasch models in health*. ISTE.

Laroche, F., Azoulay, D., Trouvin, A. P., Coste, J., & Perrot, S. (2019). Fibromyalgia in the workplace: Risk factors for sick leave are related to professional context rather than fibromyalgia characteristics-a French national survey of 955 patients. *BMC Rheumatology*, *26*(3), 44. https://doi.org/10.1186/s41927-019-0089-0

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

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, *7*(4), 328. https://www.rasch.org/rmt/rmt74m.htm

Marais, I., & Andrich, D. (2008). Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement*, *9*(3), 200–215.

Marques, A. P., Santo, A. S. D. E., Berssaneti, A. A., Matsutani, L. A., & Yuan, S. L. K. (2017). Prevalence of fibromyalgia: Literature review update. *Revista Brasileira de Reumatologia*, *57*(4), 356–363. https://doi.org/10.1016/j.rbr.2016.10.004

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

Mukhida, K., Carroll, W., & Arseneault, R. (2020). Does work have to be so painful? A review of the literature examining the effects of fibromyalgia on the working experience from the patient perspective. *Canadian Journal of Pain*, *4*(1), 268–286. https://doi.org/10.1080/24740527.2020.1820858

National Institute for Health and Care Excellence. (2021). Chronic pain (primary and secondary) in over 16s: Assessment of all chronic pain and management of chronic primary pain. Retrieved from www.nice.org.uk/guidance/ng193

National Institute for Health Research. (2023). Clinical Research Network. Retrieved from https://www.nihr.ac.uk/explore-nihr/support/clinical-research-network.htm

NHS England, & NHS Improvement (2019). Elective care high impact interventions: First contact practitioner for MSK services. Retrieved from https://www.england.nhs.uk/wp-content/uploads/2019/05/elective-care-high-impact-interventions-first-contact-practitioner-msk-services-specification.pdf

Nouri, F., Coole, C., Smyth, G., & Drummond, A. (2021). The allied health professions health and work report and the fit note: Perspectives of patients and stakeholders. *British Journal of Occupational Therapy*, *84*(2), 72–78. https://doi.org/10.1177/0308022620948763

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

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

Palstam, A., & Mannerkorpi, K. (2017). Work ability in fibromyalgia: An update in the 21st century. *Current Rheumatology Reviews*, *13*(3), 180–187. https://doi.org/10.2174/1573397113666170502152955

Rakovski, C., Zettel-Watson, L., & Rutledge, D. (2012). Association of employment and working conditions with physical and mental health symptoms for people with fibromyalgia. *Disability & Rehabilitation*, *34*(15), 1277–1283. https://doi.org/10.3109/09638288.2011.641658

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests, (Copenhagen, Danish Institute for Educational Research), with foreword and afterword by B.D. Wright*. The University of Chicago Press.

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

Reilly, M. C., Zbrozek, A. S., & Dukes, E. M. (1993). The validity and reproducibility of a work productivity and activity impairment instrument. *PharmacoEconomics*, *4*(5), 353–365. https://doi.org/10.2165/00019053-199304050-00006

Ruschak, I., Monteso-Curto, P., Rossello, L., Martin, C. A., Sanchez-Monteso, L., & Toussaint, L. (2023). Fibromyalgia syndrome pain in men and women: A scoping review. *Healthcare (Basel)*, *11*(2), 223. https://doi.org/10.3390/healthcare11020223

Salaffi, F., Di Carlo, M., Farah, S., Mariani, C., Fulginei, S., Martino, G. P., & Sarzi-Puttini, P. (2022). A cross-sectional research on female workers examining the loss of productivity caused by mild, moderate and severe fibromyalgia. *Clinical & Experimental Rheumatology*, *40*(6), 1151–1158. https://doi.org/10.55563/clinexprheumatol/hut4ft

Shemtob, L., & Asanati, K. (2021). How to get the most out of the fit note. *British Journal of General Practice*, *71*(712), 525–526. https://doi.org/10.3399/bjgp21X717665

Smith, E. V. (2002). Detecting and evaluation the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, *3*, 205–231.

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

Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: Early developments, current practice, and future challenges. *Frontiers in Rehabilitation Science*, *17*(4), 1208670. https://doi.org/10.3389/fresc.2023.1208670

Tennant, A., Penta, M., Tesio, L., Grimby, G., Thonnard, J.-L., Slade, A., Lawton, G., Simone, A., Carter, J., Lundgren-Nilsson, Å., Tripolski, M., Ring, H., Biering-Sørensen, F., Marincek, Č., Burger, H., & Phillips, S. (2004). Assessing and adjusting for cross cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: The pro-ESOR project. *Medical Care*, *42*(Suppl. 1), 37–48. https://doi.org/10.1097/01.mlr.0000103529.63132.77

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 proposed 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

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research*, *21*(4), 651–657. https://doi.org/10.1007/s11136-011-9960-1

Wilson, N., Beasley, M. J., Pope, C., Dulake, D., Moir, L. J., Hollick, R. J., & Macfarlane, G. J. (2022). UK healthcare services for people with fibromyalgia: Results from two web-based national surveys (the PACFiND study). *BMC Health Services Research*, *3*(1), 989. https://doi.org/10.1186/s12913-022-08324-4

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., Brown, T., Prior, Y., Ching, A., & Parker, J. (2023). Psychometric testing of the Rheumatoid Arthritis Work Instability Scale in employed people with fibromyalgia. *Musculoskeletal Care*, 1–13. https://doi.org/10.1002/msc.1829