



Examining Work-Related Functioning in a Physical Therapy Outpatient Clinic: Validity and Reliability of the Work Rehabilitation Questionnaire (WORQ)

Aline Husmann¹ · Reuben Escorpizo² · Monika E. Finger³

Published online: 29 August 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Purpose Musculoskeletal disorders (MSDs) are often associated with long-term sick leave, productivity loss, and reduced work functioning. However, measures that assess work-related functioning are sparse. **Objective** To assess the psychometric properties of the Work Rehabilitation Questionnaire (WORQ)-German version in patients with MSDs in an outpatient physical therapy practice. **Methods** Psychometric study including patients with MSDs with restricted work participation. Data was collected in a single physical therapy outpatient clinic. For construct validity, we developed a priori hypotheses on the correlation between the functioning part of WORQ (40 items) and other questionnaires with similar concepts. For test–retest reliability, WORQ was administered twice, 7 days apart. We examined internal consistency (Cronbach’s Alpha) and Minimal Detectable Change (MDC). Feasibility of WORQ was examined using feedback from patients and physical therapists. **Results** There were 51 study participants. Test–retest of WORQ sum score was 0.80 ($p < 0.01$) (Spearman’s rho). Internal consistency was 0.94 and MDC established at 9.2%. WORQ correlated with general health ($r = -0.49$), with HADS ($r = 0.55$), and with quality of life (WHOQOL) ($r = -0.47$). WORQ had the highest correlation with WHODAS 2.0 ($r = 0.81$). Patients rated WORQ as easy to answer and meaningful to their experience. **Conclusions** When evaluating self-reported work-related functioning, the WORQ-German version was demonstrated to be a valid, reliable, and easy to administer questionnaire for our sample of patients with MSDs in an outpatient PT clinic.

Keywords Work rehabilitation · International Classification of Functioning, Disability and Health ICF · Psychometric properties · Validity · Reliability

Introduction

Work is one of the most powerful social determinants of health; it is a central aspect of people’s lives and can influence productivity and integration in the society [1]. In contrast, being unemployed has been linked to increased mortality risk among working-age people [2, 3]. In recent years, the prevalence of chronic musculoskeletal disorders (MSDs)

has increased in many industrialized countries—leading to high burden on the individual level and a significant increase in costs [4–6].

Policy makers and stakeholders in healthcare agree that a patient-centred and holistic approach is essential to identify the needs of the patients and to successfully address work disability [7, 8]. An increasing number of multidisciplinary (including physical therapy) occupational health programmes and vocational rehabilitation (VR) measures reflect this idea, where physical therapists play a crucial role in promoting workers’ health, preventing work-related injuries, and in developing comprehensive multimodal rehabilitation programmes to improve workers’ ability to return to and stay at work [9, 10].

The role of physical therapists in occupational health and VR remains crucial to help mitigate the effects of disability on work [11]. Further, the use of the International Classification of Functioning, Disability and Health (ICF)

✉ Reuben Escorpizo
Reuben.Escorpizo@med.uvm.edu

¹ Praxis Löwencentre, Lucerne, Switzerland

² Department of Rehabilitation and Movement Science, The University of Vermont, Burlington, VT, USA

³ Department of Participation, Integration and Social Epidemiology, Swiss Paraplegic Research, Nottwil, Switzerland

as the biopsychosocial approach to plan, direct and evaluate rehabilitation outcomes is a cornerstone in physical therapists practice [12].

The ICF Core Set for Vocational Rehabilitation

Health professionals value the use of the biopsychosocial model to understand health and functioning and consider the model to be helpful in designing return-to-work (RTW) programs [13, 14]. To facilitate the use of the ICF in VR, an international, multidisciplinary group developed the ICF core sets for VR, which is a list of ICF categories that describe what a clinician (or researcher) should assess in the context of work functioning and offers guidance in the selection of work functioning questionnaires or tools [15, 16]. The ICF core set for VR was validated from the perspective of physical therapists [10]. Physical therapists expressed the need to examine work-related functioning using a reliable and valid questionnaire so they can properly evaluate the worker, predict work outcomes and implement a sound RTW strategy [17].

Work Rehabilitation Questionnaire (WORQ)

Based on the ICF core sets for VR, Finger et al. [18] developed a generic patient-reported questionnaire—the Work Rehabilitation Questionnaire (WORQ) (www.myworq.org)—which was intended to capture work-related functioning of individuals who are undergoing VR [18–20]. WORQ was originally developed in English and has an interviewer-administered and a self-reported version, and has been translated into multiple languages, including French and German using a standard cross-cultural adaptation process [19, 21].

The French and German versions of WORQ are being used in Switzerland. The psychometric properties of the French version of WORQ has been previously investigated and has shown strong test–retest reliability and a strong internal consistency in patients with MSDs who are undergoing multidisciplinary VR. The French version however was tested in the *inpatient* setting as VR is commonly delivered in the inpatient setting in Western Europe [21]. In other countries including the United States, VR is typically implemented in the outpatient setting. The German version of WORQ is being used in a physical therapy *outpatient* clinic, but its validity and reliability has not yet been examined in such a setting. Therefore, the aim of this study is to examine the psychometric properties of WORQ-German in terms of its validity, reliability, and feasibility in a physical therapy outpatient clinic.

Methods

Participants and Data Collection

The study was conducted in one outpatient physical therapy clinic in the centre of Switzerland (Lucerne).

We employed convenience sampling, with the following inclusion criteria: working age (18 to 65 years), referred to the clinic because of an MSD and have expressed to have limitations with their work, able to speak, read and write in German, and have the autonomy to make their own decision.

After the treating physical therapist confirmed the participant's eligibility criteria, the physical therapist explained the study to the prospective participant. If the patients agreed to participate in the study and provided written informed consent, they were asked to complete the “case-report form 1” (CRF1) for the first time (T1). To evaluate test–retest reliability, the participants were asked to complete “case-report form 2” (CRF2) 7 days after completing CRF1 (T2).

Case Report Form (CRF)

CRF1 contained sociodemographic information, WORQ, a question on general functioning (“Please indicate the extent of your problems in functioning in everyday life”), the single visual analogue scale of the health status questionnaire EuroQoL-5D (EQ5D) [22] (“We would like to know how good or bad your health is TODAY”), the Hospital Anxiety and Depression Scale (HADS-D) [23], the 12-item version of the WHO Disability Assessment Schedule (WHODAS 2.0) [24], the World Health Organization Quality of Life Questionnaire (WHOQoL) [25] and the Self-administered Comorbidity Questionnaire (SCQ) [26]. CRF 2 contained WORQ, two more questions on content validity and two questions on the usability of WORQ.

Instruments

WORQ consist of two parts [20]. Part one collects 17 items on relevant background information about the work situation, social support, work environment and sociodemographic data of the patient/client. Part two includes 40 questions on work-related functioning including body functions and activities and participation (Table 1). Each item is scored on a numeric rating scale from 0 to 10, (0 = no problems. 10 = complete problem). An overall summary score and four clinical sub-scores on emotion (item 4, 5, 6, 7, 8, 23), cognition (item 3, 9, 10, 17, 18, 19, 20, 24, 25, 26), dexterity (item 14, 15, 21, 22, 27, 28, 29, 34, 35, 36) and mobility (item 12, 30, 31, 32)—that have been derived from an earlier explanatory factor

Table 1 Items of the Work Rehabilitation Questionnaire (WORQ)—Part II related to work functioning

Overall in the past week, to what extent did you have problems with...	ICF	Emotion	Cognition	Dexterity	Mobility
1 ... not feeling rested and refreshed during the day?	BF				
2 ... sleeping, such as falling asleep, waking up frequently during the night or waking up too early in the morning?	BF				
3 ... remembering to do important things?	BF		✓		
4 ... your usual daily activities because you felt sad or depressed?	BF	✓			
5 ... your usual daily activities because you felt worried or anxious?	BF	✓			
6 ... being irritable?	BF	✓			
7 ... your temper?	BF	✓			
8 ... your self-confidence?	BF	✓			
9 ... thinking clearly?	BF		✓		
10 ... analysing and finding solutions to problems in day to day life?	BF		✓		
11 ... hearing?	BF				
12 ... keeping your balance while maintaining a position or during movement?	BF				✓
13 ... bodily aches or pain?	BF				
14 ... general endurance when performing physical activities?	BF			✓	
15 ... muscle strength?	BF			✓	
16 ... skin problems, such as broken skin, ulcers, bedsores and thinning of skin?	BF				
17 ... learning a new task (e.g., learning a new game, learning how to use the computer, learning how to use a tool, etc.)?	A&P		✓		
18 ... focusing attention on a specific task or e.g., filtering out distractions such as noise?	A&P		✓		
19 ... reading?	A&P		✓		
20 ... making decisions?	A&P		✓		
21 ... starting and completing a single task such as making your bed or cleaning up your desk or workplace?	A&P			✓	
22 ... carrying out your daily routine or day to day activities?	A&P			✓	
23 ... handling stress, crises, or conflict?	A&P	✓			
24 ... understanding body gestures, symbols and drawings?	A&P		✓		
25 ... starting and maintaining a conversation?	A&P		✓		
26 ... using communication devices such as using a telephone, telecommunication devices, and computers?	A&P		✓		
27 ... lifting and carrying objects weighing up to 5 kg?	A&P			✓	
28 ... lifting and carrying objects weighing more than 5 kg?	A&P			✓	
29 ... fine hand use such as handling objects, picking up, manipulating and releasing objects using the hand, fingers, and thumb?	A&P			✓	
30 ... walking a short distance (less than 1 km)?	A&P				✓
31 ... walking a long distance (more than 1 km)?	A&P				✓
32 ... moving around including crawling, climbing, and running?	A&P				✓
33 ... using transportation as a passenger?	A&P				
34 ... driving a car or any form of transportation?	A&P			✓	
35 ... getting dressed?	A&P			✓	
36 ... looking after your health such as maintaining a balanced diet, getting enough physical activity and seeing your doctor as needed?	A&P			✓	
37 ... your relationships with people?	A&P				
38 ... having sufficient money to cover your cost of living?	A&P				
39 ... seeing and recognizing an object at arm's length?	BF				
40 ... seeing and recognizing a person you know across the road (distance of about 20 meters or 66 feet)?	BF				

Components of the International Classification of Functioning, Disability and Health (ICF)

BF body function, *A&P* activities and participation)

analysis—can be calculated based on the 40 functioning items from part II [20, 27]. The clinical subscores have been developed to identify underlying patterns of functioning to support clinical decision-making and intervention allocation. Ten items are not assigned to any subscale (Table 1). The developers identified these items as being relevant to complement the picture of work-related functioning and to consider the different needs of patients with various health conditions. All WORQ scores have been confirmed by Rasch analysis [28].

EuroQoL EQ-5D EQ-5D is a standardized measurement of health status, developed by the EuroQoL-Group to offer a simple, generic measure of health for clinical and economic appraisal [29]. Consistent with other studies, we only used part two, the visual analogue scale (0 to 100) as a global indicator of General Health [30].

Hospital Anxiety and Depression Scale (HADS) HADS is commonly used to determine the levels of anxiety and depression that a patient experiences [23, 31]. HADS has 14-items on anxiety and depression. HADS has a sensitivity and specificity of about 0.80 in assessing the symptom severity of anxiety and depression disorders in patients with somatic disorders and in the general population [32].

WHO Disability Assessment Schedule (WHODAS 2.0, 12-item version) WHODAS 2.0 provides a generic standardized assessment of functioning and disability in individuals with any kind of disease, including MSDs [33, 34]. WHODAS 2.0 distinguishes well between, normal population, population with MSDs and mental disorders [35].

World Health Organization Quality of Life Questionnaires (WHOQoL) WHOQoL was developed to measure quality of life in a variety of cultural settings [36, 37]. For this study we used the five questions of WHOQoL to capture the subjective appraisal of health and well-being [38]: (1) How would you rate your quality of life? (2) How satisfied are you with your health? (3) How satisfied are you with your ability to perform your daily activities? (4) How satisfied are you with your personal relationships? (5) How satisfied are you with the conditions of your living place?

Self-administered Comorbidity Questionnaire (SCQ) SCQ is an instrument to assess comorbidity for clinical and health services research. The patient is asked about 12 medical problems, each with three scoring options [26]; (1) Do you have a problem (yes/no), (2) Do you get treatment for this problem? (3) Are you restricted in your activities?

Validity

Patients who completed CRF-1 at T1 have been included in the validity analysis.

Content Validity: Patient Perspective

Content validity of WORQ was evaluated by written questions, Q1 and Q2, at the end of CRF-2. (Q1) “From your perspective, did WORQ asks all relevant aspects concerning VR? (Yes–No)”, (Q2) “Are the answering options meaningful? (Yes–No). If no, please comment”. These questions were followed by face to face interviews with seven patients.

Construct Validity

Construct validity for the WORQ summary score was examined based on five a priori hypotheses. Our criterion to reject construct validity was that two or more hypotheses were rejected.

- (a) WORQ (problems in work-related functioning) correlates moderately negative ($r > -0.5$) with general health, measured with the EQ-5D VAS scale, where higher scores indicate better health. We know that functioning determines a major aspect of health, and we expect the same for work-related functioning [28].
- (b) WORQ correlates moderately ($r > 0.5$) and WORQ-emotion sub-score correlates highly ($r > 0.7$) with the Hospital Anxiety and Depression Scale (HADS). We found in earlier studies that emotional functioning was a critical factor in the inpatient setting [21] and expect the same for the outpatient setting.
- (c) WORQ correlates highly ($r > 0.7$) with WHODAS 2.0, because the 12 items of WHODAS 2.0 capture impaired function and disability in a similar but generic way, although not specific to the context of VR [39].
- (d) WORQ correlates weakly ($r > 0.3$) with WHOQoL, because functioning represents only a fraction of the trait of *general wellbeing* [40].
- (e) Problems in work-related functioning as assessed with WORQ correlates moderately ($r > 0.5$) with the number of comorbidities assessed with SCQ.

Feasibility: Patient and Physical Therapist Perspective

Patients answered two questions on feasibility of WORQ in CRF-2: Q3 “Did you have problems with the numeric rating scale? (Yes–No). If yes, please comment” and, Q4

“How is the length of WORQ for you? (much too long, a bit too long, good, a bit too short, much too short)”. In addition, we interviewed nine clinician physical therapists about the feasibility and use of information obtained through WORQ.

Reliability

For test–retest reliability, we choose 7 days in-between T1 and T2 to minimize recall bias and assuming no change in functioning during this period, i.e. response stability. Stability was also assessed subjectively based on a Global rating of Change scale (–5 maximum worsening to 5 maximum improvement) and moreover with a paired *t* test (parametric) or a Wilcoxon signed-rank test (non-parametric) [41]. Test–retest reliability was calculated with ICC (parametric) or Spearman rank correlation (non-parametric). Correlations measure the strength and direction of the relationship between two variables. Values for the coefficient *r* can range from 0 (no correlation) to –1 or 1 (perfect negative or perfect positive correlation); a value above 0.7 is considered as high positive [42, 43].

Internal consistency of WORQ and its clinical subscores at T1 was calculated with Cronbach’s alpha for WORQ. Cronbach’s alpha is a general coefficient of homogeneity between the items within a questionnaire. Values can range from 0 to 1 and can be interpreted as $\alpha > 0.9$ excellent, > 0.8 good, > 0.7 acceptable, > 0.6 questionable, > 0.5 poor, and $\alpha < 0.5$ —unacceptable [44].

Precision

We considered floor and ceiling effects to be present if more than 15% of participants achieved either the lowest or highest possible scores, respectively [44]. The Standard Error of Measurement (SEM) represents the smallest score change that can be interpreted as real change. SEM was calculated using Cronbach’s α as reliability coefficient R_x . $SEM = SD \sqrt{1 - R_x}$ [45, 46]. The Minimal Detectable Change (MDC), meaning the minimum amount of change in a patient’s average score that is not the result of measurement error, was calculated on the 95% probability as $MDC = 1.96 \times SEM \times \sqrt{2}$ [47].

Statistical analysis

To describe our sample, we calculated descriptive statistics and to determine normality of our data, we performed a histogram analysis and the Kolmogorov–Smirnov test. Stability was tested based on *t*-test (parametric) or related-sample Wilcoxon signed rank test (non-parametric). We

Table 2 Characteristics of study participants

Patient/client characteristics	n=51
Age (mean, years)	39.96 (12.9 SD)
Sex (male)	20 (39.2%)
Primary health condition involving	
Upper extremity	13 (25.5%)
Lower extremity	25 (49.0%)
Back	8 (15.7%)
Neck	3 (5.9%)
Endurance/muscle strength deficits	2 (3.9%)
Profession	
Light physical work	17 (33.3%)
Moderate physical work	18 (35.3%)
Heavy physical work	16 (31.4%)
Education	
Primary school or less	2 (3.8%)
Secondary school	31 (60.8%)
College or university	18 (35.4%)
Reduction in working hours:	
Off work	22 (43.1%)
< 50%	11 (21.6%)
50–100%	4 (7.8%)
Full time	14 (27.5%)
Vocational rehabilitation activities:	
Adaptation of working hours	19 (37.3%)
Workplace modification	5 (9.8%)
Work organization	6 (11.8%)
Looking for a new job	1 (1.9%)
No vocational rehabilitation activities	20 (39.2%)
Duration of sick leave	
< 3 months	29 (56.9%)
< 6 months	8 (15.7%)
< 1 year	6 (11.8%)
< 2 years	2 (3.9%)
> 2 years	6 (11.8%)
Number of co-morbidities	
0	39 (76.5%)
1	7 (13.7%)
2	3 (5.9%)
3	1 (2.0%)
4	1 (2.0%)

used Pearson correlation (parametric) or Spearman correlation (non-parametric) to test our hypothesis, depending on the distribution of data. All analysis were performed with IBM SPSS Statistics for Windows, Version 25.0 [48].

Results

In total, 51 patients completed CRF-1 and CRF-2 (Table 2). A majority of the participants was female. About 75% of the patients had problems of the upper and lower extremity, and 15% from back-related problems. Ten percent of the patients had a chronic comorbidity, e.g., neurological problem. Fifty-three percent of the patients had an accident and 47% had illness-related problems. Seventy-six percent ($n=39$) patients reported no comorbidities. Over half of the patients contacted the physical therapy clinic < 3 month after the onset of their health problem and 20% of the participants had their health problem for more than 20 months (Table 2).

Data analysis

Only 0.2% of responses on WORQ was missing. Nevertheless, missing data for WORQ were imputed using Miss Forest analysis with RStudio—a non-parametric missing value imputation for mixed-type data [49, 50]. Because histogram and statistical analysis of the Kolmogorov–Smirnov test showed that only WHODAS 2.0 and WHOQoL scores were normally distributed, we analysed our data according to non-parametric statistical methods.

Validity

Content Validity

Forty-eight of 51 (94%) patients reported that WORQ covered all relevant topics, and 47 (93%) patients found the

answer options to be meaningful, two found that the rating scale should be narrower, for example on a scale of 0 to 5 instead of 0 to 10. After completing CRF-2, seven participants reported that no valuable information on work-related functioning was missing in WORQ. They felt that WORQ facilitated a patient-centric approach to their care, because of its comprehensive set of questions allowed them to report their experience as patients. In particular, the work-related issues asked from part one of WORQ were considered valuable, although some participants found it difficult to rate the current support of their employer, given their current sick leave.

Construct Validity

Because data were not normally distributed, we used Spearman correlation to test our a priori hypothesis. (a) Consistent with our assumption WORQ correlated moderately with general health (EQ 5D) with a $r = -0.49$. (b) As hypothesized, we found a strong positive correlation between WORQ-emotion sub-score and HADS ($r = 0.71$) and moderate correlation between WORQ and HADS ($r = 0.55$). (c) As expected WORQ correlated highest with WHODAS 2.0 ($r = 0.81$) and good with the general rating of functioning scale ($r = 0.62$), both of which are measures of general functioning. (d) WORQ correlated almost moderately with quality of life measured by WHOQoL ($r = -0.47$). (e) To our surprise work-related functioning and the number of comorbidities showed no significant correlation which could partly be explained by the low number of patients (24%) who reported a comorbidity. In addition, only five participants (9.9%) reported that this comorbidity influenced their functioning (Table 3).

Table 3 Construct validity between the Work Rehabilitation Questionnaire (WORQ) and other questionnaires

Questionnaires	WORQ Average Sum score	WORQ Emotion Score	WORQ Cognition Score	WORQ Dexterity Score	WORQ Mobility Score
EQ-5D scale: “Today’s health”	-0.49**	-0.28*	-0.26	-0.38**	-0.46**
HADS					
HADS Overall score	0.55**	0.71**	0.53**	0.38**	0.13
HADS Anxiety score	0.55**	0.66**	0.43**	0.42**	0.20
HADS Depression score	0.49**	0.62**	0.54**	0.31*	0.09
WHODAS 2.0 – 12-item version	0.81**	0.53**	0.52**	0.65**	0.67**
General rating of functioning (NRS 0–10)	0.62**	0.36*	0.34*	0.32*	0.29*
WHOQoL	-0.47**	-0.35*	-0.34*	-0.33*	-0.38**
Number of comorbidities, median/range	0.10	0.10	0.04	0.11	0.08
Time off work (month)	0.22	0.22	0.30*	0.06	-0.06

EQ-5D EuroQol Instrument, HADS Hospital Anxiety and Depression Scales, WHODAS 2.0 WHO Disability Assessment Schedule 2.0 version, NRS numeric rating scale, WHOQoL World Health Organization Quality of Life Questionnaire

*Correlation is significant at the 0.05 level

**Correlation is significant at the 0.01 level

Four out of five hypothesis were confirmed. Therefore, construct validity was not rejected.

Feasibility

Forty-seven of 51 patients found WORQ useful to understand and describe their problems. The majority (78%) considered the length of WORQ appropriate, but 11 participants found WORQ to be too long. Ninety-four percent of the participants said that WORQ was easy to understand. Moreover, they found that the items of WORQ are short and simple, and easy to understand. All nine involved physical therapists reported that they have gained a significant information about how their patients with MSDs perceived their work-related functioning and their work situation. The time needed to instruct the participants on how to complete WORQ was 2–3 min. Nevertheless, the physical therapists also suggested that WORQ would be most valuable when used for patients with complex diagnosis and multiple comorbidities in the context of their restricted work participation.

Reliability

Test–Retest Reliability

Mean time between T1 and T2 was 9.1 days. WORQ sum score and sub-scores changed highly significant ($p \leq 0.001$) from T1 to T2. These results were confirmed by the patients rating of the Global rating of change scale (Median +1/

Table 5 Reliability results—internal consistency (Cronbach’s alpha) of the Work Rehabilitation Questionnaire (WORQ)

Questionnaire (n = number of items)	Cronbach’s alpha (α)
WORQ Scale (n = 40)	0.94
WORQ Subscale: emotion (n = 6)	0.91
WORQ Subscale: cognition (n = 10)	0.90
WORQ Subscale: dexterity (n = 10)	0.90
WORQ Subscale: mobility (n = 4)	0.84

Range 8 (– 3 to 5)). Test–retest reliability of WORQ was high with a Spearman’s $r = 0.79$ (Table 4).

Internal Consistency

WORQ showed excellent internal consistency with a Cronbach’s alpha of $\alpha = 0.94$. Also, the clinical subscales showed good to excellent results: emotion score $\alpha = 0.91$, cognition score $\alpha = 0.91$, dexterity score $\alpha = 0.85$ and mobility score $\alpha = 0.85$ (Table 5).

Precision of WORQ

No ceiling or floor effect was detected. The WORQ summary score ranged from 21 to 282/400 points with a median score of 79 points. The SEM was calculated as 13.53 points out of the maximal sum score of 400; and the MDC was calculated as 32.35 points what is equal to 8.09% change,

Table 4 Reliability (correlation) and questionnaire results

		Time 1	Time 2	Test–retest Spearman’s r
WORQ-40 items Average score	Median (range)	1.98 (6.53)	1.45** (6.23)	0.79
WORQ Average Emotion score	Median (range)	1.17 (9.00)	1.00** (7.83)	0.78
WORQ Average Cognition score	Median (range)	0.90 (6.90)	0.60** (6.40)	0.81
WORQ Average Dexterity score	Median (range)	3.70 (7.10)	2.70** (8.20)	0.80
WORQ Average Mobility score	Median (range)	2.75 (8.50)	2.38** (9.00)	0.86
EQ-5D scale: “Today’s health”	Median (range)	70.0 (70)		
General Functioning (NRS 0-10)	Median (range)	5.00 (8)		
WHODAS-12-item version	Median (range)	13.0 (38.00)		
WHOQoL	Median (range)	19.00 (14.00)		
HADS				
Overall score	Median (range)	8.00 (23.00)		
Anxiety score	Median (range)	4.00 (14.00)		
Depression score	Median (range)	4.00 (11.00)		
Global Change Scale (– 5 to +5)	Median (limits)		+ 1 (– 3 to 5)	

Average WORQ scores are based on functioning items from part II of WORQ, ** highly significant change from T1 to T2

EQ-5D EuroQol Instrument, HADS Hospital Anxiety and Depression Scales, WHODAS 2.0 WHO Disability Assessment Schedule 2.0 version, NRS numeric rating scale, WHOQoL World Health Organization Quality of Life Questionnaire

meaning that changes in the summary score that are higher than 32.35 points can be attributed to a real change.

Discussion

Statement of Principal Findings

This study reports on the first psychometric evaluation of work-related functioning questionnaire called WORQ in individuals with MSDs in a physical therapy outpatient clinic. Based on our findings, WORQ is a valid questionnaire to assess work-related functioning in our study sample in terms of content and construct validity, although the low median of the summary score of 79/400 points—despite the lack of floor effect—suggests that the information gain with WORQ might be higher in a population with more complex functioning problems.

WORQ showed good test–retest reliability, excellent internal consistency and established its MDC. Moreover, WORQ provided physical therapists with relevant work-related information for patients with work restriction. Patients valued WORQ as a comprehensive and easy to answer questionnaire that encouraged them to express their health-related functional problems and related work-related problems in a comprehensive and non-biomedical way. We expect that the established psychometric performance of WORQ will pave the way for consequent multi-center studies looking at the clinical utility of WORQ across disease and practice settings.

Strengths of the Study

Our study demonstrated the clinical utility of WORQ German version as an ICF-based questionnaire that assesses work-related functioning in MSDs and is easy to use in a physical therapy outpatient clinic. It is also the first study to show the benefits of WORQ in an outpatient VR situation, which is also typical in many countries outside Europe. Patients and physical therapists alike have confirmed the value of WORQ to cover the multidimensional nature of work and work-related functioning. WORQ supports physical therapists to integrate the concept of return-to-work as a participation goal into their intervention planning, amidst the focus on biomedical paradigms and impairment-focused treatment plans [51]. Our study has also provided evidence that WORQ can guide the assessment by physical therapists of functioning aspects not necessarily always considered in their clinical practice such as emotions, cognition, environmental support and relationships that may have impact on the plan of care. Another strength of the study is our fairly homogeneous sample with predominantly MSDs, which represents the typical range of health conditions referred

to an outpatient PT practice [52], and which allows—to an extent—the potential transferability of our results to other similar outpatient clinic settings.

Limitations of the Study

The weaknesses of our study include the observational nature of our study design, the convenience sampling, and the small sample size. Our convenience sample that includes participants with injury-related and illness-related MSD problems, satisfies the intent of WORQ which is to be health condition independent. Nevertheless, the mixed sample may reduce the transferability of the results to distinct populations, such as employees with work accidents [53, 54]. The small sample size could have potentially contributed to the non-normal distribution of the data. To account for the distribution and the improvement, we used nonparametric statistics, e.g., Spearman correlation to calculate test–retest reliability instead of Intraclass Correlation Coefficient (ICC). The full version of WORQ showed a high internal consistency with a Cronbach's alpha of > 0.90 , but we recognize that the alpha coefficient could be inflated because alpha is influenced by the number of items in the test [55]. Furthermore, while the study participants had significantly improved on their functioning between T1 and T2, we suspect that these changes have led to a somewhat lower reliability than our previous study in the inpatient setting. Aiming for a shorter test–retest period would have been beneficial, as change within 7 days could be expected in the outpatient practice setting. However, a time period of < 7 days can be critical due to increased recall bias [21]. Another limitation was that the data in this study was collected in one single outpatient clinic in the German-speaking region in Switzerland potentially limiting its broad applicability. Nevertheless, WORQ was cross-culturally translated to German and the language itself should not substantially impact the generalizability of WORQ to other similar settings.

Strengths and Weaknesses in Relation to Other Studies

Most questionnaires used in the context of work rehabilitating setting or occupational health other than WORQ, solely focus on work limitations [56], work participation, job satisfaction or job stability [57]. In contrast to WORQ, these questionnaires have been designed to identify persons at risk for work disability or dropout from the workforce. WORQ's patient-reported functioning, instead, when used as an initial screening questionnaire, would enable physical therapists to select specific assessment instruments to determine shared rehabilitation goals and to plan a worker-focused intervention [58]. Physical therapists may also reflect upon the

multiple aspects of functioning assessed by WORQ to define their role and expertise within a multidisciplinary team in VR. Recently a new ICF-based instrument, the Work Disability Functional Assessment Battery (WD-FAB) [59] was developed with the aim to inform a disability assessment process. Although WD-FAB uses in parts similar questions as WORQ, WD-FAB is based on computer adapted testing and both instruments serve different aims in different parts of the spectrum of work disability. WORQ, with its total of 57 items, serves as a clinical questionnaire to inform the return-to-work process and its sustainability, WD-FAB quantifies the level of disability, based on 5 to 7 adapted questions, to support the decision process for disability evaluation.

In practice, our study supported the notion that patients felt ready to discuss priorities and preference with a professional, because WORQ had helped them to elaborate their physical, psychological, emotional or work-related concerns in advance. Hence, WORQ can promote efficiency during physical therapy encounter, especially if patients have multiple comorbidities in multiple areas of functioning [60].

Furthermore, as early identification of appropriate interventions can potentially prevent acute or recurrent health problems from becoming chronic, WORQ may help to identify in the early stage rehabilitation barriers to patient's recovery and hence, may be helpful in work disability prevention.

Unanswered Questions and Needs for Future Research

WORQ-German version proved to be an instrument with good psychometric properties in the outpatient setting, and the French version and the Dutch version (separate work) reported similar results in terms of psychometric properties and usability. Therefore, we expect that the English version, as well as the other cross-culturally adapted versions, have comparable psychometrics as well, although further studies are needed in the future to verify the psychometric properties of WORQ. With our study's cross-sectional design, this study was not able to evaluate the ability of WORQ to predict return to work and did not establish its sensitivity to change. Although WORQ was designed to cover work-related functioning independent of the health condition, so far, WORQ has been evaluated in a population with predominantly MSDs. Its value remains to be proven in other health conditions, especially in mental health or combined physical and mental health conditions [10].

In conclusion, we found evidence that WORQ is a valid, reliable and easy to administer questionnaire, to evaluate self-reported work-related functioning of patients with MSDs in a physical therapy outpatient clinic. Physical therapists may use WORQ for shared decision-making and goal

setting in the context of return-to-work goals in their clinical practice. However, further studies will shed light on the utility of WORQ in diverse patient populations and settings.

Acknowledgements The authors thank all physical therapists at Loewencenter Lucerne for their help with the data collection and the patient/clients for their participation.

Author contributions AH, MF, RE—Concept/idea/research design. AH, MF, RE—Writing. AH—Data collection, AH, MF—Data analysis. AH, MF—Project management. RE—Consultation (including review of manuscript before submitting).

Funding This study was supported by the Swiss Paraplegic Research as part of a Master Thesis.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical Approval All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000 (5).

Informed Consent Informed consent was obtained from all patients for being included in the study. The Ethics Committee EKNZ, Switzerland (Project-ID: 2017-00552) granted full ethical approval of this research study.

References

1. Waddell G, Burton AK, Kendall NA. Vocational rehabilitation: what works, for whom, and when? London: The Stationary Office; 2008. <http://www.kmgph.com/assets/hwvb-vocational-rehabilitation.pdf>. Accessed 25 Jan 2018.
2. Bachmann N, Burla L, Kohler D. Gesundheit in der Schweiz – Fokus chronische Erkrankungen Nationaler Gesundheitsbericht 2015. Bern: Horgrefe Verlag; 2015.
3. Roelfs DJ, Shor E, Davidson KW, Schwartz JE. Losing life and livelihood: a systematic review and meta-analysis of unemployment and all-cause mortality. *Soc Sci Med*. 2011;72(6):840–854.
4. Selander J, Marnetoft SU, Bergroth A, Ekholm J. Return to work following vocational rehabilitation for neck, back and shoulder problems: risk factors reviewed. *Disabil Rehabil*. 2002;24(14):704–712.
5. EULAR Position Paper on Health and Safety at Work: Fighting rheumatic and musculoskeletal diseases together [database on the Internet]. 2018. https://www.eular.org/myUploadData/files/eular_position_paper_on_health_and_safety_at_work.pdf. Accessed 14 Feb 2018.
6. United States Bone and Joint Initiative (USBJI). Opportunities for action—the impact of musculoskeletal disorders on americans. Rosemont: USBJI; 2016.
7. OECD, HM. The Next Generation of Health Reforms: Ministerial Statement. Paris: 2017.
8. Nutzi M, Trezzini B, Medici L, Schwegler U. Job matching: an interdisciplinary scoping study with implications for vocational rehabilitation counseling. *Rehabil Psychol*. 2017;62(1):45–68.

9. Escorpizo R, Reneman MF, Ekholm J, Fritz J, Krupa T, Marnett SU, et al. A conceptual definition of vocational rehabilitation based on the ICF: building a shared global model. *J Occup Rehabil.* 2011;21(2):126–133.
10. Kaech Moll VM, Escorpizo R, Portmann Bergamaschi R, Finger ME. Validation of the comprehensive ICF core set for vocational rehabilitation from the perspective of physical therapists: international delphi survey. *Phys Ther.* 2016;96(8):1262–1275.
11. Boucaut R. International Federation of Physical Therapists working in Occupational Health and Ergonomics (IFPTOHE). WCPT. 2019. <https://www.wcpt.org/IFPTOHE>. Accessed 31 May 2019.
12. WCPT. The International Classification of Functioning, Disability and Health. [updated 10/2016]. <https://www.wcpt.org/icf>. Accessed 31 May 2019.
13. Escorpizo R, Stucki G. Disability evaluation, social security, and the international classification of functioning, disability and health: the time is now. *J Occup Environ Med.* 2013;55(6):644–651.
14. Escorpizo R, Finger ME, Glassel A, Gradinger F, Luckenkemper M, Cieza A. A systematic review of functioning in vocational rehabilitation using the International Classification of Functioning, Disability and Health. *J Occup Rehabil.* 2011;21(2):134–146.
15. Selb M, Escorpizo R, Kostanjsek N, Stucki G, Ustun B, Cieza A. A guide on how to develop an international classification of functioning, disability and health core set. *Eur J Phys Rehabil Med.* 2015;51(1):105–117.
16. Finger ME, Escorpizo R, Glassel A, Gmunder HP, Luckenkemper M, Chan C, et al. ICF Core Set for vocational rehabilitation: results of an international consensus conference. *Disabil Rehabil.* 2012;34(5):429–438.
17. Chamberlain MA, Fialka Moser V, Schuldt Ekholm K, O'Connor RJ, Hecceg M, Ekholm J. Vocational rehabilitation: an educational review. *J Rehabil Med.* 2009;41(11):856–869.
18. Finger ME, Escorpizo R, Bostan C, De Bie R. Work Rehabilitation Questionnaire (WORQ): development and preliminary psychometric evidence of an ICF-based questionnaire for vocational rehabilitation. *J Occup Rehabil.* 2014;24(3):498–510.
19. Finger M, De Bie RA, Nowak D, Escorpizo R. Development and Testing of an ICF-based Questionnaire to evaluate functioning in vocational rehabilitation: The Work Rehabilitation Questionnaire (WORQ). In: Escorpizo R, Brage S, Homa DB, Stucki G, editors. *Handbook of vocational rehabilitation and disability evaluation*. Basel: Springer; 2015.
20. Finger M, Escorpizo R. Homepage: Work Rehabilitation Questionnaire (WORQ). Swiss Paraplegic Research/ICF Research Branch. 2013. [updated December 2017]. <http://myworq.org/>. Accessed 19 March 2019.
21. Finger ME, Wicki-Roten V, Leger B, Escorpizo R. Cross-Cultural Adaptation of the Work Rehabilitation Questionnaire (WORQ) to French: a valid and reliable instrument to assess work functioning. *J Occup Rehabil.* 2018;29(2):350–360.
22. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199–208.
23. Snaith RP. The Hospital Anxiety And Depression Scale. *Health Qual Life Outcomes.* 2003;1:29.
24. World Health Organization (WHO). WHODAS-II - disability assessment schedule training manual: a guide to administration. Geneva: WHO; 2004.
25. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Qual Life Res.* 2004;13(2):299–310.
26. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis Rheum.* 2003;49(2):156–163.
27. Finger M, Schwegler U, Escorpizo R. Poster presentation: psychometric properties and factorial structure of the ICF based Work Rehabilitation Questionnaire (WORQ)—French Version. *Work Disability Prevention & Integration*; Amsterdam 2016.
28. Finger ME, Escorpizo R, Tennant A. Measuring work-related functioning using the self-reported Work Rehabilitation Questionnaire (WORQ): metric properties based on Rasch analysis. *Int J Environ Res Public Health.* 2019;16(15):E2795. <https://doi.org/10.3390/ijerph16152795>.
29. Solberg TK, Olsen JA, Ingebrigtsen T, Hofoss D, Nygaard OP. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J.* 2005;14(10):1000–1007.
30. Kirchberger I, Cieza A, Biering-Sorensen F, Baumberger M, Charlifue S, Post MW, et al. ICF Core Sets for individuals with spinal cord injury in the early post-acute context. *Spinal Cord.* 2010;48(4):297–304.
31. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* 1983;67(6):361–370.
32. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res.* 2002;52(2):69–77.
33. Üstün BT, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ.* 2010;88:815–823. <https://doi.org/10.2471/BLT.09.067231>.
34. Posl M, Cieza A, Stucki G. Psychometric properties of the WHO-DAS II in rehabilitation patients. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation.* 2007;16(9):1521–1531.
35. Andrews G, Kemp A, Sunderland M, Von Korff M, Ustun TB. Normative data for the 12 item WHO Disability Assessment Schedule 2.0. *PLoS ONE.* 2009;4(12):e8343.
36. Conrad I, Matschinger H, Riedel-Heller S, von Gottberg C, Kilian R. The psychometric properties of the German version of the WHOQOL-OLD in the German population aged 60 and older. *Health Qual Life Outcomes.* 2014;12:105. <https://doi.org/10.1186/s12955-014-0105-4>.
37. Power M, Quinn K, Schmidt S. Development of the WHOQOL-old module. *Qual Life Res.* 2005;14(10):2197–2214.
38. Glassel A, Finger ME, Cieza A, Treitler C, Coenen M, Escorpizo R. Vocational rehabilitation from the client's perspective using the International Classification of Functioning, Disability and Health (ICF) as a reference. *J Occup Rehabil.* 2011;21(2):167–178.
39. Pösl M, Cieza A, Stucki G. Psychometric properties of WHODAS II in rehabilitation patients. *Qual Life Res.* 2007;16(9):1521–1531.
40. Renwick R, Nourhaghighi N, Manns PJ, Rudman DL. Quality of life for people with physical disabilities: a new instrument. *Int J Rehabil Res.* 2003;26(4):279–287.
41. Field A. *Discovering statistics using SPSS*. London: SAGE Publications; 2009.
42. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: a Practical Guide*. Cambridge: Cambridge University Press; 2011.
43. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24(3):69–71.
44. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42.
45. Brown JD. Standard error vs. Standard error off measurement. *JALT Testing & Evaluation SIG Newsletter* [Internet].

- 1999;3(1):20–25. <http://hosted.jalt.org/test/PDF/Brown4.pdf>. Accessed 14 Aug 2017.
46. Tighe J, McManus IC, Dewhurst NG, Chis L, Mucklow J. The standard error of measurement is a more appropriate measure of quality for postgraduate medical assessments than is reliability: an analysis of MRCP(UK) examinations. *BMC Med Educ.* 2010;10:40. <https://doi.org/10.1186/1472-6920-10-40>.
47. Donoghue D, Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *J Rehabil Med.* 2009;41(5):343–346.
48. IBM Corp. IBM SPSS Statistics for Window, Version 25.0. Armonk, NY: IBM Corp; Released 2017.
49. RStudio Team. RStudio: Integrated development environment for R. Boston: RStudio, Inc.; 2015. <http://www.rstudio.com/>. Accessed 30 Aug 2017.
50. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics.* 2012;28(1):112–118.
51. Stenner R, Palmer S, Hammond R. What matters most to people in musculoskeletal physiotherapy consultations? A qualitative study. *Musculoskelet Sci Pract.* 2018;35:84–89. <https://doi.org/10.1016/j.msksp.2018.03.005>.
52. Childs JD, Harman JS, Rodeghero JR, Horn M, George SZ. Implications of practice setting on clinical outcomes and efficiency of care in the delivery of physical therapy services. *J Orthop Sports Phys Ther.* 2014;44(12):955–963.
53. Hou WH, Tsao JY, Lin CH, Liang HW, Du CL. Worker's compensation and return-to-work following orthopaedic injury to extremities. *J Rehabil Med.* 2008;40(6):440–445.
54. Bartys S, Frederiksen P, Bendix T, Burton K. System influences on work disability due to low back pain: an international evidence synthesis. *Health Policy.* 2017;121(8):903–912.
55. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>.
56. Lerner D, Amick BC III, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. *Med Care.* 2001;39(1):72–85.
57. Revicki D, Ganguli A, Kimel M, Roy S, Chen N, Safikhani S, et al. Reliability and Validity of the Work Instability Scale for Rheumatoid Arthritis. *Value Health.* 2015;18(8):1008–1015.
58. American Physical Therapy Association (APTA). Physical therapist examination and evaluation: Focus on tests and measures. *Guide to Physical Therapist Practice 3.0.* 2014.
59. Marfeo EE, Haley SM, Jette AM, Eisen SV, Ni P, Bogusz K, et al. Conceptual foundation for measures of physical function and behavioral health function for Social Security work disability evaluation. *Arch Phys Med Rehabil.* 2013;94(9):1645–1652 (e1642).
60. Licher S, Heshmatollah A, van der Willik KD, Stricker BHC, Ruiter R, de Roos EW, et al. Lifetime risk and multimorbidity of non-communicable diseases and disease-free life expectancy in the general population: a population-based cohort study. *PLoS Med.* 2019;16(2):e1002741.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.