



# Cross-Cultural Adaptation of the Work Rehabilitation Questionnaire (WORQ) to French: A Valid and Reliable Instrument to Assess Work Functioning

Monika E. Finger<sup>1,2,3</sup> · Virginie Wicki-Roten<sup>1,5</sup> · Bertrand Leger<sup>1,5</sup> · Reuben Escorpizo<sup>1,2,4</sup>

Published online: 26 June 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

**Purpose** The Work Rehabilitation Questionnaire (WORQ) is a patient-reported instrument to assess work related functioning in vocational rehabilitation (VR) and work, based on the International Classification of Functioning, Disability and Health (ICF) core set for VR. The objective of this study was to cross-culturally adapt WORQ to French and to evaluate its psychometric properties. **Methods** The cross-cultural adaptation followed a dual-panel approach. Psychometrics was examined in one VR-centre in the French speaking part of Switzerland. Test–retest reliability was analyzed with intraclass correlation coefficient (ICC), internal consistency was assessed with Cronbach’s alpha. Construct validity was determined by convergence to the self-reported general functioning scale and the Hospital Anxiety and Depression Scales (HADS). The association of patient’s return-to-work expectation within 6-months and having a case manager was examined. The usability of WORQ-French was tested in 10 VR patients. **Results** Eighty-nine patients with musculoskeletal injuries were included. WORQ-French showed excellent internal consistency (0.968) and a high test–retest reliability (0.935). WORQ-French was positively associated with self-reported general functioning ( $r=0.662$ ) and both HADS scales ( $r=0.56–0.57$ ). Neither the patient’s return-to-work expectation nor having a case manager were significantly correlated with WORQ-French. Usability in terms of understandability of questions and response options was found to be good. Seven patients rated the length of WORQ-French as good, while two found the instrument a little too long and one found it too long. **Conclusions** WORQ French is a valid, reliable, and easy to administer instrument to assess self-reported work functioning given our study setting and sample characteristics.

**Keywords** Vocational rehabilitation · Work · WORQ · Functioning · ICF · Psychometrics

## Introduction

Vocational rehabilitation (VR) is an integral part of a disability management strategy and is essential to support workers and persons with disabilities willing to work to “secure, retain and advance in suitable gainful work participation”. Such a process will, in turn, enhance the workers’ integration or reintegration into the society [1]. VR interventions not only encourage the person’s participation in daily living and the society but also reduce the days of sick leave and help to prevent health-related impairments from becoming chronic [2]. Individuals who benefit from VR not only have a primary health condition but also suffer from various comorbidities, leading to a complex level of disability [3–5]. Moreover, overall problems in functioning are assumed to contribute to a decreased or lost workability [6, 7], where workability has been described as “the

✉ Monika E. Finger  
monika.finger@paraplegie.ch

<sup>1</sup> Swiss Paraplegic Research (SPF), Nottwil, Switzerland

<sup>2</sup> ICF Research Branch, a Cooperation Partner Within the WHO Collaborating Centre for the Family of International Classifications in Germany (at DIMDI), Nottwil, Switzerland

<sup>3</sup> Caphri Research School, Maastricht University, Maastricht, The Netherlands

<sup>4</sup> Department of Rehabilitation and Movement Science, University of Vermont, Burlington, VT, USA

<sup>5</sup> Department of Research and Clinical Quality Control, Clinique Romande de Réadaptation, Sion, Switzerland

balance of the workers' resources and work demands". In this context resources or deficits refer to "functional abilities, professional skills and professional values" [8]. Mitigating measures to address work disability is critical to ensure that the work functioning of an individual is restored to the optimal level possible.

For VR to be effective, a carefully planned and comprehensive vocational assessment that addresses all relevant factors is crucial. Identifying the individual who needs the appropriate return-to-work support, optimizing the timing and selecting the best possible VR interventions requires a comprehensive understanding of the complex nature of the relationship between work and health resources [9–11].

Previous studies showed that the *International Classification of Functioning, Disability and Health (ICF)*, an integrative biopsychosocial model of the World Health Organization (WHO), provides a suitable reference framework to address the complexity and multi-faceted nature of VR [12, 13]. ICF core sets, an abbreviated list of relevant domains based on the ICF, have been developed over the years to make the application of the ICF to specific settings in health care in a way that is practical and easy [14]. The categories from the ICF core set for VR were selected systematically and empirically and were intended to describe work-related functioning independent of a health condition, the respective VR setting or a specific time point throughout the return-to-work continuum.

While the ICF core sets can provide the basis of what domains of functioning can be assessed, it does not necessarily provide a way on how such an assessment can be undertaken. Hence, the search for a patient-centered measure that can be used to assess the ICF core set for VR has been made. No such instrument was found thus, during our search, efforts were made to develop a new questionnaire. The first version of the Work Rehabilitation Questionnaire (WORQ) is an interviewer-administered. WORQ consists of two parts: part 1 includes sociodemographic information, environmental factors, and work-related information and part two contains items on work-related functioning consisting of body functions and activities and participation domains. The development process of WORQ was described in detail elsewhere [15]. Feasibility tests of WORQ showed high patient satisfaction with the questionnaire. However, the interviewing health and work professionals raised the issue of whether or not a patient-reported version of WORQ can also be available to enhance feasibility and ease of use [16, 17]. To address the lack of such an instrument, a self-reported version of WORQ was developed. In the self-reported version, the changes mainly concern the introductory text, a shift from three support items from part two to the first part and the wording of items related to personality and temperament functions. (<http://www.mywork.org>).

Switzerland has four official national languages. The majority of the population (63%) speak German, hence the first version of WORQ was made available in German. However, the data collection for the purpose of testing the psychometric s is still ongoing. However, the second largest language group in Switzerland speaks French (23% of the population) [18]. Because there is no French version of WORQ to ensure applicability of the questionnaire in the majority of the Swiss population, a study was undertaken to develop and examine the psychometric properties of the French version of WORQ. Hence, our study-aim was to translate and cross-culturally adapt the patient-reported version of WORQ to French and to evaluate the fundamental psychometric properties of WORQ-French in a sample of persons undergoing vocational rehabilitation in a specialized center in Switzerland.

## Methods

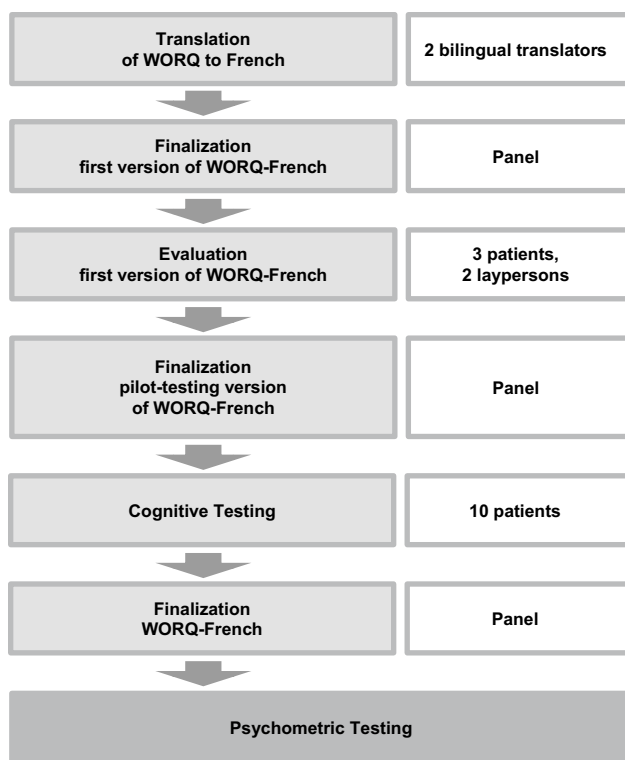
This single-center observational psychometrics study was approved by the ethical committee of the canton Valais in Switzerland (CCVEM 005/15) and conducted according to the principles outlined in the Declaration of Helsinki. All patients signed a written informed consent form before participating in the study.

### WORQ Instrument

Part one of WORQ contained items on sociodemographic information and work-related items including age, gender, education, profession, work status, current VR interventions and items about family support, the support provided by the superior/supervisor, and the labor system support. Part two of WORQ has 40 functioning items, each with a response scale of 0–100 (0 = No problem to 100 = Complete problem). The internal consistency of the interviewer-administered version of WORQ showed Cronbach's alpha of 0.887 and test–retest agreement of 0.789 (Spearman correlation) [15].

### Step 1: Translation and Cross-Cultural Adaptation of WORQ

The translation and cross-cultural adaptation was based on guidelines proposed by Beaton et al. and then later modified to follow a dual-panel approach [19, 20]. In this approach, the English-language version of WORQ was forward-translated into French in Switzerland by a bilingual panel including, two experts in questionnaire development and validation, three bilingual experts (two Swiss-Francophone, one French), one VR specialist and two of the developers (Fig. 1). This first version of WORQ-French was then evaluated by two lay persons (non-patient) as well as three patients



**Fig. 1** Process of cross-cultural adaptation of WORQ to French

undergoing VR interventions due to a shoulder problem, a hip and a knee injury [21, 22]. These five individuals completed the questionnaire and were asked to comment about the understandability of WORQ. The panel discussed the findings and, if needed, the members agreed on adaptations. At this point, the WORQ modification as the result of this process underwent pilot testing in the next stage. The next stage consists of cognitively testing of WORQ by a group of ten patients. These patients were diverse regarding age, gender, diagnosis, and education. In the cognitive testing, the patients were asked to complete the pilot-testing version of WORQ and to provide feedback on the clarity and understandability of the items. They were also asked the ease of completing WORQ. Finally, the bilingual panel verified the second version of WORQ based on the information from the pretest and finalized the French version of WORQ (<http://www.myworq.org>).

## Step 2: Psychometric Evaluation of WORQ-French Version

### Participants and Procedures

The psychometric evaluation took place in a Swiss rehabilitation teaching hospital (Clinique Romande de Réadaptation, Switzerland). Participants with musculoskeletal

injuries were referred from all of the French-speaking countries of Switzerland, which includes urban and industrial city centers or more rural regions. Inclusion period lasted from March 2015 to February 2016. Patients were contacted by a research assistant and invited to participate if they were (a) aged between 18 and 65 years, (b) were participating in a vocational rehabilitation intervention due to musculoskeletal condition, and (c) is proficient in French (oral and written language) to be able to complete questionnaires. After giving informed written consent, the patient was provided with a patient case report form CRF-T0 including WORQ.

### Variables and Instruments

A patient case report form (CRF-T0), along with WORQ, was completed by the participant upon admission into the VR program in the clinic. The CRF-T0 collected information pertaining to sociodemographics, such as age, gender, and family status. The CRF-T0 also contained a self-evaluation of general health measure using a numeric rating scale from 0 to 10 (0 = excellent and 10 = poor), a self-evaluation of general functioning using a numeric rating scale from 0 to 10 (0 = no problem and 10 = extensive problem), as well as an appraisal of the current state of health using the single item of the EQ-5D (a vertical visual analogue scale from 0 to 100 with the bottom anchor as “the worst state of health imaginable” and the top anchor as “best state of health imaginable”). These variables have been proven to be valuable indicators in understanding the health situation of persons in the context of vocational rehabilitation [5, 23]. In addition, the rehabilitation center provided further data that was collected at admission in the context of routine clinical practice. This data included injury location (upper extremity, lower extremity, trunk/back and polytrauma), having a case manager (response option: yes/no), and “case severity”. “Case severity” was determined by the treating physician and based on the Abbreviated Injury Scale (AIS) [24]. The AIS is an anatomical-based coding system created by the Association for the Advancement of Automotive Medicine to classify and describe the severity of injuries. Finally a variable reflecting the patient’s expectation if he or she will be back at work within 6 month (response option: yes/no) was reported [25]. The patients expectation was found to be a strong predictor for the resumption of work in patients with subacute and chronic musculoskeletal conditions [26–28]. The *Hospital Anxiety and Depression Scale (HADS)* was used to detect anxiety and depression, variables that are firmly related to rehabilitation outcome and successful return to work [29–31]. The *HADS* is a 14-item scale that was initially developed by Zigmond and Snaith [32] and is commonly used to determine the levels of anxiety and depression that a patient is experiencing. It was initially created to detect the perception of anxiety and depression in

people with physical health restrictions. Seven of the items relate to anxiety and seven relate to depression. A subscore of 8–10 is considered as borderline and a score of 11+ is considered as anxiety or depression disorder.

### Usability

Usability of WORQ-French was tested in a group of ten patients that participated in the psychometric testing. Understandability of the items in WORQ was addressed by asking “Did you have any difficulties understanding these questions? If YES, please write down which question/s”. For suitability of response options, the question was: “Did the response options make sense to you? Please comment”. For feasibility and appropriateness of length of the entire questionnaire, the participants were asked to choose one of the following options: “too long”, “a little long”, “a good length”, “a little short”, and “too short”.

### Reliability

**Internal Consistency** Internal consistency of WORQ-French was examined based on data from the full study sample at admission, using Cronbach’s alpha. Cronbach’s alpha is a general coefficient of homogeneity between items. Values for the coefficient  $\alpha$  can range from 0 (no internal consistency) to 1 (perfect internal consistency). Coefficients above 0.75 are considered moderate, above 0.75 they are regarded as good, and excellent above 0.9 [33].

**Test–Retest Reliability** The first 50 study participants were invited to complete WORQ-French for a second time 7 days after their initial completion to evaluate test–retest reliability. An average score was calculated for each participant by summing up the scores from the 40 equally-weighted items from part two, divided by the total number of items answered. As a single exception, item 34 “Overall in the past week, to what extent did you have problems with driving a car or any form of transportation?” had an additional answer option, “not applicable”. This option related to the situation of a person without driver’s license or without an option to drive because of a lack of car. In this case, a problem with driving would not relate to a disability. In the case of a “not applicable” the item was excluded from the sumscore calculation, i.e., for a total of 39 items. Test–retest reliability was calculated based on the average score using interclass correlation  $ICC_{2,1}$  [34].

### Precision

Floor and ceiling effects were considered to be present if more than 15% of participants achieved either the lowest or highest possible scores, respectively [35].

The standard error of measurement (SEM) was calculated to assess response stability, meaning that the SEM is the amount of error that can be considered as measurement error. SEM was calculated using Cronbach’s  $\alpha$  as reliability coefficient  $R_x$ .  $SEM = SD \sqrt{1 - R_x}$  [36, 37]. 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  $MDC = 1.96 \times SEM \times \sqrt{2}$  [38].

### Validity

Content validity examines the extent to which the domain of interest, work-related functioning, in this study is comprehensively covered by the items in the questionnaire. We examined content validity in interviews by asking patients to comment on the comprehensiveness of WORQ-French in relation to their specific situation in VR.

Taking into account the theoretical and conceptual framework of the ICF based on which WORQ was developed, WORQ evaluates work-related functioning by taking into account, amongst other things, mental, emotional, physical and movement-related body functions and activities. Therefore a positive moderate association of “work-related functioning” as measured by WORQ with *self-evaluated general functioning* representing the encompassing concept of functioning is assumed. Anxiety and depression as measured by the two dimensions of HADS are expected to correlate only with items in WORQ that address mood and emotion, hence, what we could assume to be a moderate correlation of the HADS scores with the WORQ sumscore in our musculo-skeletal population. As functioning is seen as a determinant of health in the ICF conceptually, we assume that the *self-evaluated general health* rating, and the *overall health* (VAS 0–100) will also be positively associated and reasonably so given the characteristics of these also overarching global concepts. Finally, we expect that patients with better functioning abilities (WORQ) are more likely to return-to-work (patient’s expectation of return-to-work within 6 months) than patients with lower functioning. In contrast, we suspected that those participants, who were supported by a case manager, showed more significant problems in work functioning than those without because case management is typically provided in more complex and severe cases.

### Statistical Analysis

Descriptive statistics were used to describe the sample. Distribution of data (normality) was tested based on histogram analysis and Kolmogorov–Smirnov and Shapiro–Wilk test [39] to determine if Spearman correlation or Pearson correlation should be used. 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 highly positive [40]. Imputation for missing data in WORQ items was done with RStudio using MissForest—non-parametric missing value imputation for mixed-type data [41, 42]. As predictor variables, we used the WORQ variables supported by the following variables: sex, age, *self-evaluation of general functioning* and *self-evaluation of general health*. All calculations were performed using the software package IBM SPSS Statistics for Windows, Version 24.0, released 2016 [43].

## Results

### Step One: Translation and Cross-Cultural Adaptation

The bilingual panel experienced no significant problems in finding consensus on the translation of the functioning items in part two of WORQ regarding meaning and style. A discussion on the anchor definition of the scale led to a translation of “Aucun Problème” for “No problem” and “Problème grave” for “Complete problem”, as the most appropriate expressions. The cross-cultural adaptation of part one (socio-demographic and work-related items) revealed the need to take into account items as they hold relevance to local and regional characteristics and settings, such as school systems and availability of VR services in items 6, 7, and 14 (<http://www.myworq.org>).

In the first evaluation of WORQ-French, the wording of the items revealed only minor word changes, such as item 12 in part one, where the word “paramedical” was changed to “therapeutic” and in part two where the verb “bouger” (to move) was replaced by the verb “se déplacer”. In contrast to the wording of the items, all three patients, and one lay person reported serious problems with the visual-analog scale (0–100) that was used to evaluate the functioning items in part two. Therefore, the panel and the developers of WORQ decided to use the numeric rating scale (0–10) instead but maintaining the same anchor definitions of 0 = No problem to 10 = Complete problem. The participants understood the numeric rating scale of 0–10 better, and they felt to have a better estimation of how they feel about the item. Two patients reported some problems with answering items from part one related to their work status, work situation and vocational rehabilitation interventions but all patient reported good understandability of the functioning items in part two of WORQ. Only one person raised the issue of the numeric rating scale to be not meaningful in completing the items. Further, after two minor changes in the wording of item 5: “Which of the following describes your current work status best?” and item 14: “What kind of work or vocational intervention are you receiving now?” in part one, a final

version of WORQ-French was approved for psychometric testing.

### Step Two: Psychometric Evaluation

#### Participants

Eighty-nine patients completed CRF-T0. Ten patient out of the 89 provided feedback on the usability of WORQ after its completion.

Fifty consecutive participants out of the initial 89 also completed the patient record form “CRF-T1”, which was administered 7 days after the CRF-T0 for test–retest. The CRF-T1 contained WORQ-French, the *self-evaluated general functioning* rating, and the *self-evaluated general health* rating. For participant characteristics, they were predominantly male and married. Although over 25% of the participants were still at work prior to admission to the clinic, the remaining participants were on average off work for more than 10 month. A majority of the participants suffered from an upper extremity injury (44%). Around 40% of the participants took part in more than one vocational intervention, and 26–34% were actively looking for a new job while also 30–32% were engaged in activities to maintain their current job.

With an average WORQ sumscore of 3.4/10 (n = 50) and 4.5/10 (n = 89), patients rated to have fewer problems in work-related functioning compared to *general functioning* rating as 5.22/10 (n = 50) and 5.56/10 (n = 89) in the whole population. In the HADS depression and anxiety scales is a score between 8 and 10 considered as “borderline case”. This indicates that our study population, with an anxiety score of 9 or 8.8 respectively, was burdened with a substantial amount of anxiety (Table 1).

#### Usability

Participants rated the usability of WORQ from fair to very good. However, three work-related items from part one were flagged by the participants. Item 8: “What is your current job or profession or if currently not working, what is the last job or profession you worked in (job title)?”, item 9: “What kind of business, industry or service is (or was) your job in?” and item 10: “What kind of work are (or were) you doing?” were found to be misleading and overlapping by three persons. These three also complained that they had to formulate the answers themselves and that WORQ didn’t provide simple response options to choose from. Especially defining specific work tasks in item 10 was considered as difficult and somewhat redundant to the job type asked for in question 8. All patients reported good understandability of the functioning items in part two and no problems with the NRS (0–10). Appropriateness of length of the entire questionnaire was



**Table 1** Characteristics of respondents, study 1

Characteristics	Test–retest sample n = 50 (%)	Validation sample n = 89 (%)
Age, mean, years	42.6 (10.3 SD)	44.0 (41.7–46.4 SD)
Sex, male	45 (90.1%)	82 (92.1%)
Family status		
Single	9 (18%)	15 (16.9%)
Married/cohabitant	25 (50%)	51 (57.3%)
Separated/divorced	16 (32%)	22 (24.8%)
Widowed	0	1 (1.1%)
Education		
Less than primary school	3 (6%)	5 (5.6%)
Primary school	11 (22%)	16 (18.0%)
Secondary school	22 (44%)	46 (51.7%)
College or university	13 (26%)	18 (20.2%)
Post graduate education	1 (2%)	4 (4.5%)
Working contract (yes)	20 (40%)	30 (33.7%)
Work status (working)	14 (28%)	23 (25.8%)
Time off work if not working (in weeks)	53.5 (33.1 SD) mean 50.0 (117 range) median	55.4 (37.2 SD) mean 43.0 (193.5 range) median
VR activities		
Work skill training	10 (20%)	27 (30.4%)
Apprenticeship or stage	2 (4%)	3 (3.4%)
Activities to maintain job	16 (32%)	27 (30.3%)
Looking for a new job	13 (26%)	30 (33.7%)
Work capacity evaluation	9 (18%)	19 (21.4%)
Physical training	18 (36%)	30 (33.6%)
Injury location		
Upper extremity	24 (48%)	39 (43.8%)
Lower extremity	15 (30%)	37 (41.6%)
Trunk/back	7 (14%)	10 (11.2%)
Poly-trauma	4 (8%)	3 (3.4%)
General health	5.0 (2.5 SD)	4.8 (2.4 SD)
General functioning	5.2 (2.3 SD)	5.6 (2.1 SD)
WORQ average score	3.4 (1.97 SD)	4.5 (1.9 SD)
HAD anxiety	9.0 (3.8 SD)	8.8 (4.0 SD)
HAD depression	7.2 (4.3 SD)	6.7 (4.0 SD)

rated as “a good length” by seven participants, two rated it as “a little too long” and one person rated it as “too long”.

### Reliability

The psychometric testing of WORQ-SR revealed excellent internal consistency (Cronbach’s  $\alpha = 0.968$ ), although the high internal consistency may be influenced by the high number of items and potential redundancy amongst the items. The test–retest reliability was also high with an  $ICC_{2,1} = 0.935$  (CI 0.889–0.963). These findings indicate the reliability of WORQ for use as a single measure to evaluate functioning in the context of work. Data quality of WORQ was good with randomly missing values of only 2.3% in total

or 0–4 (3.56%) missing values per item, which were consequently imputed. Seven persons answered “not applicable” in item 34 “problems with driving” resulting in an average score out of 39 items. However, no change in test–retest reliability or construct validity was found when “not applicable” cases were eliminated via sensitivity analysis.

### Precision of WORQ

No ceiling or bottom effect was detected. The WORQ sum-scores ranged from 6 to 346/400 points.

The SEM was calculated as 0.323 points out of the maximal average score of 10; and the MDC was calculated as 0.895 points, meaning that changes in the average-score that

are higher than 0.895 can be attributed to a real change. This value will be helpful in determining change or the stability of WORQ when collected over multiple time points.

## Validity

**Content Validity** In the context of the content evaluation, all patients found that WORQ covered all relevant aspects of work-related functioning, although one person missed items on “off work activities”, such as household chores, sport and community activities. This person considered the degree of problems that he experienced off work as relevant, as it impacted on his work-life balance. He thought that having a considerable amount of problems off-work would lead to a decrease in work functioning.

**Construct Validity** Correlation of WORQ-French with other standard instruments can be seen in Table 2. As expected, a higher WORQ-French score was moderately associated with a higher score on self-reported *general functioning* (Pearson correlation = 0.662). WORQ-French also correlated moderately with both HADS scales, reflecting psychological distress that is represented in WORQ with six items related to the construct of “mood”, which highlights the contribution of emotional functioning to the overall work-related functioning in our population of mostly chronically injured workers.

Self-perceived health measured with the *self-evaluated general health* rating and the *overall health* also correlated significantly with the WORQ-French score. The *self-evaluated general health* rating correlated positively with WORQ, however, as expected due to the opposing direction of its visual analog scale, with the lower anchor “worst health possible” and the top anchor “best imaginable health”, the *overall health score* correlated negatively with WORQ-French.

Contrary to our expectations, neither the patient’s expectation of return-to-work within 6 months nor having a case manager (in severe cases) showed any significant correlation

with WORQ-French (work related functioning) or *self-reported general functioning*. This may indicate that patients based their expectations factors unrelated to their problems in functioning and more related to personal factors or environmental factors such as work itself and work environment [26, 27].

## Discussion

In this study, we performed a cross-culturally adaption of WORQ into French and evaluated its fundamental psychometric properties. Our findings suggest that WORQ-French is a reliable, valid, and easy-to-use instrument to assess self-reported work-related functioning. WORQ is a useful instrument to describe an individual’s work functioning using a biopsychosocial framework, while at the same time recognizing the influence of the environment. WORQ can also assess work-specific and general functioning aspects in the context of VR and support the planning of work participation strategies for people with various health conditions.

WORQ-French represents the first psychometric evidence on the self-reported version of WORQ. A dual-panel approach was successfully used for the cross-cultural adaptation as recommended in current guidelines [19, 44]. The adaptation of the second (functioning) part of WORQ resulted in few problems, neither concerning the language nor the cultural setting. In contrast to the second part, in the first part items six “When thinking about your work or vocational rehabilitation program: Are you currently”, seven “What is the highest level of education that you have completed?” and fourteen “What kind of work or vocational intervention are you receiving now?” that are related to school systems and occupational training or VR interventions, gave rise to extensive discussions. The region-specific naming of the respective school level, such as secondary school or real school instead of high school, or finding the appropriate examples to describe setting or system specific VR interventions have been identified to be crucial for the participants to provide exact and reliable answers in the self-reported instrument. The need of context-specific adaptation of these work and education related items was further confirmed by two physicians from France, who were asked by the authors to comment on WORQ-French. In further cross-cultural adaptations, a specific focus has to be set on context-specific items from part one to achieve cultural appropriateness without losing comparability across countries [45].

One primary concern for the use of WORQ in clinical practice, but even more so in research, is the length of WORQ. The notion of the global perspective taken by WORQ is leading to 40 functioning items, of which not all may appeal to everyone. Nevertheless, because of its broadness, WORQ-French is excellently capable of capturing the

**Table 2** Construct validity of WORQ-French

	WORQ sumscore
General functioning	0.662*** Pearson correlation
HADS anxiety score	0.564*** Pearson correlation
HADS depression score	0.570*** Pearson correlation
General health	0.484*** Pearson correlation
Current state of health	−0.552*** Pearson correlation
Patient expectation	
RTW within 6 month	−0.017 Spearman correlation
Case management	0.108 Spearman correlation

\*\*\*The relationship of WORQ sumscore with the corresponding variable is significant at  $p=0.01$

diversity of functioning problems that may be caused by an increasing number of co-morbidities found in individuals in VR [4, 46]. For example, a 42 years old construction worker is referred to VR to evaluate future work perspectives, after a severe work accident. He is not able to return to his former work due to severe lower extremity injuries. Besides, he suffers from of a previous mild traumatic brain injury with minor attentional deficits that didn't bother him in his old job. He controlled his diabetes II well with sport and diet before the accident, but due to a lack of movement diabetic complications, such as ulcers at his coccyx have appeared. In such a situation WORQ-French can provide a fast and comprehensive overview as needed for a patient-centered assessment, to provide the necessary interventions promptly and to inform decisions concerning sustainable work in the future [16]. Nevertheless to encounter the challenge of the length of WORQ, a brief version of part two, containing a subset of 13 items was developed. These 13 items represent the body function, and activity categories from the brief ICF core set for VR complemented with the categories from the generic set [14, 47]. The WORQ-brief assess critical aspects of work-related functioning, such as “energy and drive”, “emotional functions”, “cognitive functions” and mobility (<http://www.myworq.com>). To which extent WORQ-brief fulfills the needs for a screening instrument has to be further evaluated and is currently tested in the English version of WORQ-brief.

A finding that had a retroactive effect on the overall development of WORQ was the apparent confusion around the use of the VAS (0–100) by the patients in the first test. In the first, interviewer-administered version of WORQ, the VAS was well accepted by the interviewers as well as by the patients. Nevertheless, the change from the VAS to an NRS (0–10) in the self-reported version of WORQ resulted in a high patient satisfaction and an increased reliability compared to the interviewer-administered VAS version [48]. This finding is also supported by the literature, concluding that the NRS had better compliance rates and responsiveness and were reported to be user-friendly relative to VAS [49–51]. As a result of this study, the developers of WORQ decided to replace the VAS scale with NRS 0–10 in the revised interviewer-administered version of WORQ, as well as in all future self-reported versions of WORQ-SR.

When comparing reliability and validity of WORQ-French to the initial interviewer-administered version of WORQ, reliability substantially improved [15]. The change of scale may attribute to this improvement. Another reason could be the fact that in the first study test–retest was assessed over a period of 14 days whereas in the current study the test–retest period was 7 days, what may be better to evaluate the clinical stability of an instrument. With a 9% minimal detectable change (MDC), WORQ-French shows a good precision, not only when considering the

heterogeneous population evaluated in this study, but also when comparing WORQ-French to other patient-reported measures (PROM) such as the Disabilities of the Arm, Shoulder, and Hand (DASH) with and MDC of 17% or the Oxford Shoulder Score (OSS) with an MDC of 10% [52]. These promising results support the use of WORQ-SR as a reliable and valid way to evaluate patient-reported functioning in the context of return to work and employment.

WORQ was designed to assess the functioning of individuals in VR independent of the health condition, throughout the whole continuum of the return-to-work process. WORQ was developed from the ICF core set for VR. This concept-based approach assumes per se concept validity, what was supported in this study by the patients. Nevertheless, no leisure time activities are addressed in WORQ what is due to the experts of the ICF core set consensus conference, where interdisciplinary VR experts explicitly decided to concentrate on work-related functioning. This decision may be questioned, considering the ongoing debate on the importance of work-life balance. Nevertheless, the 40 functioning items of the self-reported version of WORQ provide information on body functions and activities, such as lifting, walking and relating to others that are also relevant to leisure time activities and may be used as proxy measures for related activities, such as sport and other hobbies [53, 54].

Similar to the interviewer-administered version of WORQ, the items of part two of the self-reported version, allow creating a functioning profile. Such a profile can be used to evaluate abilities and resources or potential, and identify areas of problems for an individual and may then serve as a basis for intervention and case management planning [16, 23].

As work-related functioning measured by WORQ-French can be assumed to be subarea of *general functioning*, WORQ-French showed a good correlation with general functioning. The higher score assessed with the *general functioning rating* compared to the WORQ average sum-score in our study may be due to the fact that the WORQ average score is assessed based on 40 tangible items, related to emotional, cognitive, physical and social aspects of functioning. Depending on the location and type of injury it can be expected that the patients score high on items related to their condition and low on items that are not affected by the respective health condition. This may lead to an average score that is lower than the overall functioning rated on one single 0–10 rating scale, because the patient is focusing on their problems and not on their resources or unaffected functional abilities. Nevertheless the correlation coefficient of 0.66 may indicate that some items in WORQ may weigh more than others in the patient's overall rating of functioning, and these items may be different from patient to patient. In addition, general functioning was assessed with only one NRS 0–10. It can be assumed that the participants



included their recent experiences with activities, their “off-work” situation as well as their current well-being in their rating [55, 56]. A good correlation of WORQ-French with its six items on mood with the HADS scores confirmed that mental and psychological functioning play a significant role in self-perceived functioning in the context of vocational rehabilitation and employment, what is consistent with the literature [57, 58]. Mood-related and cognitive in addition to movement-related aspects of functioning are also found to be most relevant in non-musculoskeletal health conditions such as heart diseases, cancer and neurologic diseases such as multiple sclerosis or traumatic brain injury. These findings may contribute to the external validity of WORQ in non-musculoskeletal conditions [59–62].

To our astonishment, the individual’s expectation to return-to-work within 6 months showed no correlation with functioning (WORQ and general functioning) at all. Although we know from the literature that the ability of patients to predict improvement of functioning is limited, we assumed that patient would base their return-to-work prediction on their current functioning status [63], what apparently was not the case in our mostly chronic population. That patients do not support their decision with their level of work-related functioning is also in line with the fact that “having a case manager”, a proxy for injury severity, did also not significantly correlate with functioning. We presume that mainly environmental factors, such as having a work contract, support of the superior and type of work influenced the decisions of the study participants. Nevertheless, the influence of work-related functioning on the prediction of return-to-work is not well studied and has to be further evaluated [64, 65].

Another issue in the context of VR is the question if self-reported functioning creates a reliable picture of a person’s abilities or if a clinician-rated assessment such as functional capacity evaluation may be needed. Current research suggests that our ability to predict return-to-work can be improved by combining self-reported and clinical information on functioning [66]. As clinical testing of work functioning is time-consuming, expensive and relies on skilled health professionals, to use a reliable and valid self-reported instrument such as WORQ, as a screening instrument may help to identify the appropriate clinical evaluation and targeted intervention in a timely and inexpensive way. Anyway, in light of patient-centred care and the person-centeredness of healthcare services, it is indispensable to evaluate and integrate the patient’s perception of his or her problems, as easily covered by WORQ-SR [67–69].

Despite the convincing results of the psychometric evaluation of WORQ-French, we caution the interpretation and generalization of our results because in this study we employed convenience sampling of patients with predominantly traumatic musculoskeletal injuries from a single VR

center. Although our sample represents the general population at the VR department of the study center, the fact that < 10% of the participants included in this study were females could limit the transferability of results to settings of different gender proportion. The cross-sectional data for the construct validation cannot provide any basis for causality. These results have to be confirmed in other clinical settings and patient groups and with diverse health conditions, to ensure the external validity of WORQ-French. Besides, ongoing longitudinal studies will further evaluate assumed underlying factors of the self-reported version of WORQ-French, as well as its predictive value on return to work.

In conclusion, we found evidence that WORQ-French is a valid, reliable and easy to administer instrument to evaluate self-reported work-related functioning given our study setting and sample characteristics. Results of WORQ may be used to guide intervention planning and document changes in functioning throughout the VR process. However, further studies will shed light on the use of WORQ in clinical practice and research, as well as in diverse patient populations and settings.

**Acknowledgements** The authors would like to thank Olivier Dériaz, Dominique Monnin and Carolina Fellinghauer for their help with the cross-cultural adaptation. We would also like to thank Wolfgang Segerer for providing technical consultation and Christina Ehrmann for statistical support. This project was funded by a research grant from the Swiss Accident Insurance Company (suva).

## References

1. International Labour Organization, editor. ILO vocational rehabilitation and employment (disabled persons) convention (no. 159) and recommendation (no. 168): United Nations convention on the rights of persons with disabilities. Geneva: ILO; 2008.
2. Seing I, MacEachen E, Stahl C, Ekberg K. Early-return-to-work in the context of an intensification of working life and changing employment relationships. *J Occup Rehabil.* 2015;25(1):74–85.
3. Marzolini S, Oh PI, Alter D, Stewart DE, Grace SL. Musculoskeletal comorbidities in cardiac patients: prevalence, predictors, and health services utilization. *Arch Phys Med Rehabil.* 2012;93(5):856–862.
4. Ervasti J, Vahtera J, Pentti J, Oksanen T, Ahola K, Kivekas T, et al. The role of psychiatric, cardiometabolic, and musculoskeletal comorbidity in the recurrence of depression-related work disability. *Depress Anxiety.* 2014;31(9):796–803.
5. Finger ME, Glassel A, Erhart P, Gradinger F, Klipstein A, Rivier G, et al. Identification of relevant ICF categories in vocational rehabilitation: a cross sectional study evaluating the clinical perspective. *J Occup Rehabil.* 2011;21(2):156–166.
6. Schmidt C, Bernert S, Spyra K. [Concerning the impact of psychological comorbidity for chronic back pain: frequency, reduced earning capacity pension and rehabilitation aftercare in the course of the rehabilitation cohorts 2002–2009]. *Rehabilitation (Stuttg).* 2014;53(6):384–389.
7. Catalina-Romero C, Ruilope LM, Sanchez-Chaparro MA, Valdivielso P, Cabrera-Sierra M, Fernandez-Labandera C, et al. Factors influencing return-to-work after cerebrovascular disease: the

- importance of previous cardiovascular risk. *Eur J Prev Cardiol.* 2015;22(9):1220–1227.
8. Hasselhorn MH. Work ability—concept and assessment. In: Weber VM, Warr N, editors. *Achieving business excellence—health, well-being and performance.* Bielefeld: Bertelsmann Stiftung, BKK Bundesverband; 2008.
  9. Power P. *A guide to vocational assessment.* 5 ed. Austin: Pro-Ed; 2013.
  10. Breeding RR. The utility of proactive vocational assessment in advancing consumer empowerment. *Vocat Eval Career Assess Prof J.* 2005;2(2):39–56.
  11. Homa D, DeLambo D. Vocational assessment and job placement: chapter 8. In: Escorpizo R, Brage S, Homa D, Stucki G, editors. *Handbook of vocational rehabilitation and disability evaluation.* New York: Springer; 2015. pp. 161–187.
  12. World Health Organization (WHO). *International classification of functioning, disability and health.* Geneva: WHO; 2001.
  13. Saltychev M, Kinnunen A, Laimi K. Vocational rehabilitation evaluation and the international classification of functioning, disability, and health (ICF). *J Occup Rehabil.* 2013;23(1):106–114.
  14. 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.
  15. 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.
  16. Finger ME, Selb M, De Bie R, Escorpizo R. Using the international classification of functioning, disability and health in physiotherapy in multidisciplinary vocational rehabilitation: a case study of low back pain. *Physiother Res Int.* 2015;20(4):231–241.
  17. Portmann Bergamaschi R, Escorpizo R, Staubli S, Finger ME. Content validity of the work rehabilitation questionnaire-self-report version WORQ-SELF in a subgroup of spinal cord injury patients. *Spinal Cord.* 2014;52(3):225–230.
  18. Swiss Federal Statistical Office. Languages declared as main languages, 2015. <https://www.bfs.admin.ch/bfs/en/home/statistics/population.assetdetail.2260545.html>. Accessed 21 Nov 2017.
  19. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186–3191.
  20. Hagell P, Hedin PJ, Meads DM, Nyberg L, McKenna SP. Effects of method of translation of patient-reported health outcome questionnaires: a randomized study of the translation of the rheumatoid arthritis quality of life (RAQoL) instrument for Sweden. *Value Health.* 2010;13(4):424–430.
  21. World Health Organization (WHO). *WHODAS-II—disability assessment schedule training manual: a guide to administration.* Geneva: WHO; 2004.
  22. O'Connor J, McCaughan D, McDauid C, Booth A, Fayter D, Rodriguez-Lopez R, et al. Orthotic management of instability of the knee related to neuromuscular and central nervous system disorders: systematic review, qualitative study, survey and costing analysis. *Health Technol Assess.* 2016;20(55):1–262.
  23. 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.
  24. Haasper C, Junge M, Ernstberger A, Brehme H, Hannawald L, Langer C, et al. [The Abbreviated Injury Scale (AIS). Options and problems in application]. *Unfallchirurg.* 2010;113(5):366–372.
  25. Iakova M, Ballabeni P, Erhart P, Seichert N, Luthi F, Deriaz O. Self perceptions as predictors for return to work 2 years after rehabilitation in orthopedic trauma inpatients. *J Occup Rehabil.* 2012;22(4):532–540.
  26. Rashid M, Kristofferzon ML, Nilsson A, Heiden M. Factors associated with return to work among people on work absence due to long-term neck or back pain: a narrative systematic review. *BMJ Open.* 2017;7(6):e014939.
  27. Steenstra IA, Munhall C, Irvin E, Oranye N, Passmore S, Van Eerd D, et al. Systematic review of prognostic factors for return to work in workers with sub acute and chronic low back pain. *J Occup Rehabil.* 2017;27(3):369–381.
  28. Opsahl J, Eriksen HR, Tveito TH. Do expectancies of return to work and Job satisfaction predict actual return to work in workers with long lasting LBP? *BMC Musculoskelet Disord.* 2016;17(1):481. <https://doi.org/10.1186/s12891-016-1314-2>.
  29. O'Donnell ML, Varker T, Holmes AC, Ellen S, Wade D, Creamer M, et al. Disability after injury: the cumulative burden of physical and mental health. *J Clin Psychiatry.* 2013;74(2):e137–e143.
  30. Carriere JS, Thibault P, Sullivan MJ. The mediating role of recovery expectancies on the relation between depression and return-to-work. *J Occup Rehabil.* 2015;25(2):348–356.
  31. Ferdiana A, Post MW, Finger M, Bultmann U, Escorpizo R. Assessment of health-related quality of life of clients in vocational rehabilitation: association with depressive symptoms and type of services. *Eur J Phys Rehabil Med.* 2014;50(3):343–353.
  32. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* 1983;67(6):361–370.
  33. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–163.
  34. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19(1):231–240.
  35. 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.
  36. Brown JD. Standard error vs. Standard error off measurement. *JALT Testing & Evaluation SIG Newsletter [Internet].* 1999;3(1). <http://hosted.jalt.org/test/PDF/Brown4.pdf>. Accessed 14 Dec 2017.
  37. 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(1):40. <https://doi.org/10.1186/1472-6920-10-40>.
  38. 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.
  39. Field A. Exploring assumptions. In: Field A, editor. *Discovering statistics using SPSS.* 3 ed. London: SAGE; 2009. pp. 131–196.
  40. George D, Mallery P. Bivariate correlation. In: George D, Mallery P, editors. *SPSS for windows 23 step by step: a simple guide and reference.* 14 ed. New York: Taylor & Francis; 2016. pp. 139–148.
  41. RStudio (2015). *RT. RStudio: Integrated development environment for R.* RStudio, Inc. Boston, MA. <http://www.rstudio.com/>. Accessed 15 Oct 2016.
  42. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics.* 2012;28(1):112–118.
  43. IBM SPSS Statistics for Windows. 24.0 ed: IBM; 2016.
  44. Epstein J, Osborne RH, Elsworth GR, Beaton DE, Guillemin F. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *J Clin Epidemiol.* 2015;68(4):360–369.
  45. Harkness J, Bilgen I, Córdova Cazar A, Hu M, Huang L, Lee S, et al. Questionnaire design. Guidelines for best practice in cross-cultural surveys. <http://www.ccsq.isr.umich.edu/>. Accessed 15 Sept 2016.

46. Boot CR, Hogg-Johnson S, Bultmann U, Amick BC 3rd, van der Beek AJ. Differences in predictors for return to work following musculoskeletal injury between workers with and without somatic comorbidities. *Int Arch Occup Environ Health*. 2014;87(8):871–879.
47. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun BT, Stucki G. Identification of candidate categories of the international classification of functioning disability and health (ICF) for a generic ICF core set based on regression modelling. *BMC Med Res Methodol*. 2006;6(1):36. <https://doi.org/10.1186/1471-2288-6-36>.
48. Voutilainen A, Pitkaaho T, Kvist T, Vehvilainen-Julkunen K. How to ask about patient satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. *J Adv Nurs*. 2016;72(4):946–957.
49. Rausch M, Zehetleitner M. A comparison between a visual analogue scale and a four point scale as measures of conscious experience of motion. *Conscious Cogn*. 2014;28:126–140.
50. Martins PC, Couto TE, Gama AC. Auditory-perceptual evaluation of the degree of vocal deviation: correlation between the Visual Analogue Scale and Numerical Scale. *Codas*. 2015;27(3):279–284.
51. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manag*. 2011;41(6):1073–1093.
52. van Kampen DA, Willems WJ, van Beers LW, Castelein RM, Scholtes VA, Terwee CB. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). *J Orthop Surg Res*. 2013;8(1):40. <https://doi.org/10.1186/1749-799X-8-40>.
53. Jakobsen K, Lillefjell M. Factors promoting a successful return to work: from an employer and employee perspective. *Scand J Occup Ther*. 2014;21(1):48–57.
54. Proper KI, van den Heuvel SG, De Vroome EM, Hildebrandt VH, Van der Beek AJ. Dose-response relation between physical activity and sick leave. *Br J Sports Med*. 2006;40(2):173–178.
55. Huijnen IP, Verbunt JA, Peters ML, Seelen HA. Is physical functioning influenced by activity-related pain prediction and fear of movement in patients with subacute low back pain? *Eur J Pain*. 2010;14(6):661–666.
56. Wand BM, Chiffelle LA, O'Connell NE, McAuley JH, Desouza LH. Self-reported assessment of disability and performance-based assessment of disability are influenced by different patient characteristics in acute low back pain. *Eur Spine J*. 2010;19(4):633–640.
57. Tankumpuan T, Utriyaprasit K, Chayaput P, Itthimathin P. Predictors of physical functioning in postoperative brain tumor patients. *J Neurosci Nurs*. 2015;47(1):E11–E21.
58. Rasmussen MU, Amris K, Rydahl-Hansen S, Danneskiold-Samsøe B, Mortensen EL, Christensen R, et al. Are the changes in observed functioning after multi-disciplinary rehabilitation of patients with fibromyalgia associated with changes in pain self-efficacy? *Disabil Rehabil*. 2017;39(17):1744–1752.
59. Giovannetti AM, Schiavolin S, Brenna G, Brambilla L, Confalonieri P, Cortese F, et al. Cognitive function alone is a poor predictor of health-related quality of life in employed patients with MS: results from a cross-sectional study. *Clin Neuropsychol*. 2016;30(2):201–215.
60. Sun Y, Shigaki CL, Armer JM. Return to work among breast cancer survivors: a literature review. *Support Care Cancer*. 2017;25(3):709–718.
61. Bush EJ, Hux K, Guetterman TC, McKelvey M. The diverse vocational experiences of five individuals returning to work after severe brain injury: a qualitative inquiry. *Brain Inj*. 2016;30(4):1–15.
62. Waszkowska M, Szymczak W. Return to work after myocardial infarction: a retrospective study. *Int J Occup Med Environ Health*. 2009;22(4):373–381.
63. Nieuwenhuijsen K, Cornelius LR, de Boer MR, Groothoff JW, Frings-Dresen MHW, van der Klink JLL, et al. Predicting improvement of functioning in disability claimants. *J Occup Rehabil*. 2014;24(3):410–418.
64. de Vries G, Koeter MW, Nieuwenhuijsen K, Hees HL, Schene AH. Predictors of impaired work functioning in employees with major depression in remission. *J Affect Disord*. 2015;185:180–187.
65. Cancelliere C, Kristman VL, Cassidy JD, Hincapie CA, Cote P, Boyle E, et al. Systematic review of return to work after mild traumatic brain injury: results of the international collaboration on mild traumatic brain injury prognosis. *Arch Phys Med Rehabil*. 2014;95(3 Suppl):S201–S209.
66. Kuijer PP, Gouttebarga V, Wind H, van Duivenbooden C, Sluiter JK, Frings-Dresen MH. Prognostic value of self-reported work ability and performance-based lifting tests for sustainable return to work among construction workers. *Scand J Work Environ Health*. 2012;38(6):600–603.
67. Levesque JF, Haggerty J, Beninguisse G, Burge F, Gass D, Beaulieu MD, et al. Mapping the coverage of attributes in validated instruments that evaluate primary healthcare from the patient perspective. *BMC Fam Pract*. 2012;13(1):20. <https://doi.org/10.1186/1471-2296-13-20>.
68. Howell D, Liu G. Can routine collection of patient reported outcome data actually improve person-centered health? *Healthc Pap*. 2011;11(4):42–47 (**discussion 55–48**).
69. Gagnier JJ. Patient reported outcomes in orthopaedics. *J Orthop Res*. 2017;35(10):2098–2108.