J Occup Rehabil (2018) 28:377–390
DOI 10.1007/s10926-017-9729-7

Reliability and Validity of the Work and Well-Being Inventory (WBI) for Employees

A. A. Vendrig[1] · **F. G. Schaafsma**[2]

Published online: 8 September 2017

Abstract *Purpose* The purpose of this study is to measure the psychometric properties of the Work and Wellbeing Inventory (WBI) (in Dutch: VAR-2), a screening tool that is used within occupational health care and rehabilitation. Our research question focused on the reliability and validity of this inventory. *Methods* Over the years seven different samples of workers, patients and sick listed workers varying in size between 89 and 912 participants (total: 2514), were used to measure the test–retest reliability, the internal consistency, the construct and concurrent validity, and the criterion and predictive validity. *Results* The 13 scales displayed good internal consistency and test–retest reliability. The constructive validity of the WBI could clearly be demonstrated in both patients and healthy workers. Confirmative factor analyses revealed a CFI>.90 for all scales. The depression scale predicted future work absenteeism (>6 weeks) because of a common mental disorder in healthy workers. The job strain scale and the illness behavior scale predicted long term absenteeism (>3 months) in workers with short-term absenteeism. The illness behavior scale moderately predicted return to work in rehab patients attending an intensive multidisciplinary program. *Conclusions* The WBI is a valid and reliable tool for occupational health practitioners to screen for risk factors for prolonged or future sickness absence. With this tool they will have reliable indications for further advice and interventions to restore the work ability.

✉ A. A. Vendrig
lexvendrig@gmail.com

[1] Vledder Medical Centre, Pastorieweg 16, 8381 AX Vledder, The Netherlands

[2] Department of Public and Occupational Health, VU University Medical Centre, EMGO+ Institute for Health and Care Research, PO Box 7057, 1007 MB Amsterdam, The Netherlands

Keywords Screening tool · Return to work · Occupational health care · Predictor variables

Introduction

The Work and Well-being Inventory (WBI) (in Dutch: VAR-2) is a Dutch instrument originally developed in the 90' to assess potential obstacles for work resumption in chronic pain patients attending a rehabilitation program. At the time it was common in rehabilitation medicine to assess depression and other psychiatric conditions pre-program. However, clinical experience revealed more and more that work-related factors such as job dissatisfaction and feelings of being misunderstood by the employer were as important to understand why sometimes a successful rehabilitation did not result in a successful return to work. Over the years, the evidence for the contribution of psychosocial factors to the return to work in chronic pain patients expanded, and now these factors are considered the most important predictors of work absenteeism and return to work in both musculoskeletal and common mental disorders [1–3]. This growth in knowledge and understanding of recovery and work ability have also been incorporated in the further development of the WBI. The WBI is a comprehensive and practical screening tool for practitioners in occupational health care. It provides an overview of the most important risk factors for (un) successful return to work. The WBI combines a prediction of the risk of future absenteeism in employees and return to work of patients suffering musculoskeletal and common mental disorders, together with an assessment of the underlying factors that contribute to that risk. Since its introduction in the 90's there has been a growing use of the WBI in the fields of occupational medicine and rehabilitation in the

Netherlands, and therefore we want to present an overview of the conducted research with the WBI.

Conceptual Background

The WBI has been entirely developed from a clinician's perspective. The main question that guided the development of the instrument was: what are the topics that may contribute to a prolonged absenteeism? What topics does an occupational physician want to scrutinize in a new patient? The topics are conceptualized as risk factors or 'causes' of prolonged absenteeism, or as support or buffer for the impact of stressors to prevent prolonged absenteeism. The more these risk factors or support factors are present, both in term of amount and intensity, the more complex the patient is for the occupational health practitioner. There is growing research evidence that workplace factors influence disability outcomes [4]. The model of the WBI resembles to a certain extent the job demands-resources (JD-R) model of occupational stress [5]. The difference between the WBI-model and the JD-R model, is that the JD-R focusses largely on occupational stress, whereas the WBI includes the person's home life and personal characteristics as well, and is largely based on a disability paradigm. Figure 1 shows the WBI model (the five domains of the WBI are shown in bold).

The main assumption of the model is that an imbalance between stressors and support leads to mental health symptoms. These symptoms, together with a negative disability perception (illness behavior), prolongs the absenteeism, irrespective of the reason for the absenteeism. The WBI considers stressors, support and person variables as indirect influences on absenteeism. The WBI also considers personality and coping to have a direct effect on symptoms, and an indirect effect by creating new stressors. For example, perfectionism may create symptoms 'in the head' because of the strain the person puts on him or herself, but perfectionism may also create stressors in the 'real world' because, for example, the person engages into too much tasks at work, and this may in return produce stress symptoms.

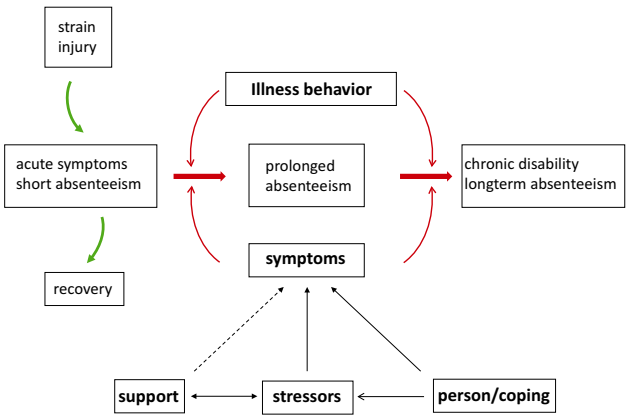

Fig. 1 The model of the WBI

Features of the WBI

The original WBI item set was the result of 1 year of gathering typical utterances of patients about their life, work and symptoms. The items are formulated as short and simple statements. The WBI has 82 items distributed among 13 scales. These 13 scales are divided among five domains. These domains are: support, stressors, personal characteristics, symptoms, and subjective disability or illness behavior. Figure 2. shows an overview of the domains and scales of the WBI. The response categories for the symptoms are: 'not', 'sometimes', 'often' and 'most.' The respondent is asked to judge how often each symptom has bothered him or her during the last week. The response categories for the

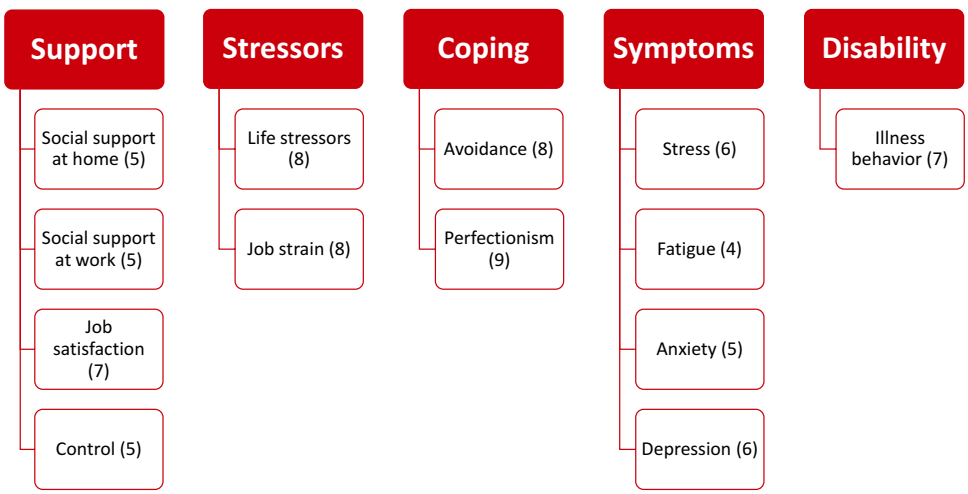

Fig. 2 Overview of the five domains and 13 scales[1] of the WBI

[1] Between brackets are the number of items for each scale.

other items are: 'not', 'somewhat', 'mostly' and 'fully'. The respondent has to judge each item to which extent he or she agrees with it.

A summary of the 13 scales is provided in the appendix. Each scale has five categories: low, below average, average, above average and high, corresponding with Dutch occupational guidelines to make a distinction between disorders (high), symptoms but not a disorder (above average), normal (average and below) [6].

The aim of this paper is to present an overview of the reliability and validity if the WBI.

Methods

Subjects

Table 1 shows a summary of the samples used in this paper. All samples contain complete data sets. Patients of sample C received a clinically established DSM-IV or DSM-V diagnosis [7]. The diagnosis was established by a psychologist or psychiatrist. The diagnosis distress or burnout (sample D) was established by an occupational physician using the distress/burnout decision tree that is included in the Dutch multidisciplinary guideline for distress and burnout [8]. Subjects of sample F and G participated in a 4-week intensive multidisciplinary treatment program for chronic pain. They all suffered long standing disabling (nonspecific) back pain complaints and were off work mostly for 6 months or longer. Before entering the program, an orthopedic surgeon had ruled out specific somatic pathology of the spine. A specific cause of the fatigue was ruled out by a rehabilitation physician in patients with fatigue complaints.

Analyses

Reliability

Internal Consistency and Test–Retest Reliability The internal consistency of the 13 WBI scales was determined with Cronbach's alpha for working employees (samples A, B), and for patients with a diagnosis of mental disorder (samples C, D). The threshold for deciding whether or not the scale's internal stability is sufficient was set at .70, whereby alpha's between .70 and .80 are regarded as acceptable, alpha's between .80 and .90 as good, and alpha's above .90 as excellent [9]. Alpha's for the smaller constructs (e.g. symptoms) should be higher than alpha's for the broader constructs (e.g. personality characteristics). Very high reliabilities (.95 or higher) are not desirable, as this indicates that the items may be redundant. The stability (test–retest reliability) of each scale was established in a subsample (n = 88) of sample A by calculating correlations (Pearson r) between the two test administrations for each scale. The time interval was about 2 weeks.

Validity

Construct Validity As a first step in discerning the construct validity item-rest correlations were calculated. We hypothesized the item-rest correlations to surpass the correlations between the item and all other scales. To date, an item should share more variance with its own scale than with any other scale. The procedure was carried out for the normal sample and patient sample separately.

The structure of the WBI was previously established in the development phase with an explorative factor analysis conducted separately for each domain. However, for instruments containing multiple domains such as the WBI, there is the risk of an enmeshment of relationships between items (e.g., items measuring job strain and fatigue may be strongly related although they go with distinct constructs) and shared variance of items measuring the same construct. From a psychometric perspective the WBI is more a battery of questionnaires measuring five distinct, although related domains, than being a solitary instrument. For this paper we tested the supposed factor structure using other samples than the sample used when the original instrument was developed.

We performed a confirmative factor analysis (CFA) and tested a 4-factor model for the support domain, a 2-factor

Table 1 Summary of study sample characteristics

Dataset	Description of sample	Size (n)	% Female	Mean age (SD)
A	Employees of different companies	912	53	42.4 (11.0)
B	Employees of a home care organization participating in a health survey	388	74	42.2 (11.4)
C	Psychiatry outpatients	98	61	40.4 (10.6)
D	Distressed/burnout patients (occupational health service)	69	54	43.5 (9.5)
E	Employees (working and sick leave) (occupational health service)	112	55	44.3 (9.6)
F	Rehabilitation patients (chronic back pain)	353	52	43.9 (10.0)
G	Rehabilitation patients (chronic fatigue complaints)	450	58	43.8 (10.0)
H	Employees, sick leave (occupational health service)	132	66	41.1 (10.4)

model for the stressors domain, a 2-factor model for the coping domain, a 4-factor model for the symptoms domain, and a 1-factor model for the disability domain.

The comparative fit index (CFI), Tucker–Lewis Index (TLI) and Root Mean Square Error of Approximation (RMSEA) were used to evaluate the fit of the model. A CFI and TLI of >.90 are generally considered as an adequate fit although >.95 seems better [10]. For good models the RMSEA should be <.06 [10]. The CFA was performed for the employees sample (A) and patients sample (C, D) separately to see whether the factor structure would be the same for samples with varying degrees of psychosocial problems. For the 2-factor domains a CFI was calculated for three models: a 1-factor model, a 2-factor model and a 2-factor bi-factor model. For the 4-factor domains a CFI was calculated for three models: a 1-factor model, a 4-factor model and a 4-factor bi-factor model.

Next, we regressed each scale onto the other scales and calculated the standardized Beta coefficients and the explained portion of variance (R^2). Since Cronbach's alpha is an estimate of the proportion of the total variance of a scale that is error free, and R^2 is the proportion of the total variance that the scale shares with the other scales, the difference between alpha and R^2 is an estimate of the unique variance of the scale [11]. Since the samples of employees and patients have quite different sample sizes, and multiple variables (12 scales) were included in the model, we used the adjusted R^2 statistic. The regression analyses were performed for the employees sample (A) and patients sample (C, D) separately.

Finally, we calculated zero-order and partial correlations between the domains. First, the raw scores of the scales were transformed to z-scores, and the z-scores of the domain's scales were summed. Second, correlations were performed between the domains. Third, partial correlations were performed to detect the unique variance shared by each set of two domains. The correlations were calculated for the employees sample (A) and patients sample (C, D) separately. We hypothesized low correlations between the support—and the other domains, high correlations between the disability—and the symptoms domain, and moderate correlations for the remaining associations between domains. Finally, we calculated Fisher's Z to test the equality of the partial correlations in the employees and patients samples. We hypothesized the aforementioned relationships between domains to be similar among varying types of health problems (employees versus patients).

Content Validity

When the WBI was developed several procedures were adopted to achieve content validity: theories guided the scale construction and item formulation, and the items were based on prototypal statements (narratives) of subjects. Next, the items were judged by independent experienced clinicians, and ambiguous items were dropped.

Concurrent Validity

Concurrent validity refers to the expected high correlations between related measurements. Correlations (Pearson r) were calculated between the WBI scales and the following instruments:

(a) *Checklist Individual Strength* (CIS20R). The CIS20R [12] consists of 20 items and was used to measure fatigue complaints. The CIS20R has been validated in the Dutch working population [13]. The CIS20R was administered in sample D (all subjects).
(b) *Fear-Avoidance Beliefs Questionnaire (FABQ)*. The work subscale (FABQ-W; seven items) of the Dutch translation [14] of the FABQ [15] was used. The FABQ focuses specifically on how a patient's fear-avoidance beliefs about physical activity and work may affect and contribute to his/her low back pain. The FABQ has 16 questions scaled from 0 to 6. The FABQ had been filled in by 63 subjects of sample F.
(c) *Four-Dimensional Symptom Questionnaire (4DSQ)*. The 4DSQ [11] was used to measure anxiety and depression. The 4DSQ encompasses 50 items in total. We used the anxiety (12 items) and depression (six items) scales. The 4DSQ is a valid self-report questionnaire to measure distress, depression, anxiety and somatization in primary care patients [11]. 44 subjects of sample D filled in the 4DSQ.
(d) *Job Content Questionnaire* (JCQ). The Dutch version [16] of the JCQ was used for the measurement of social support at work, control and job strain. We used the Co-worker support (four items) and Supervisor support (four items) scales for the assessment of social support at work; the Psychological demands scale (five items) was used to measure job strain; the Decision authority scale (three items) was used to measure Control. The JCQ has been developed to test the job-demand-control-support (JDC-S) model and has dominated occupational research on occupational stress [17]. 66 subjects of sample A filled in the JCQ.
(e) *Need for Recovery after Work Scale*. The NRW scale [18] measures work-related fatigue and consists of 11 yes/no items. The Need for Recovery after Work Scale was found to be a reliable and valid indicator of work-related fatigue [19]. The NRW scale was administered in sample D (all subjects).
(f) *Social Readjustment Rating Scale* (SRRS). The SRRS [20] was used to provide a standardized measure of the impact of a wide range of common stressors. Sub-

jects are asked to tally a list of 43 life events based on a relative score. The stressor at the top (death of a spouse) receives 100 points, and the stressor lowest in rank (minor violation of law) receives 11 points. Gerst et al. [21] tested the reliability of the SRRS, and found that the rank ordering of life events remained extremely consistent. The SRRS measures at an ordinal level resulting in a disproportional weight of extreme values and a skewed distribution. We therefore applied a log transformation to the data. 66 subjects of sample A filled in the SRRS.

(g) *Utrecht's Coping List* (UCL). We used Dutch version [22] of the UCL [23] and used the avoidance (eight items) and active tackling (seven items) scales to measure an avoidant coping style. 66 subjects of sample A filled in the UCL.

(h) *Work Ability Index* (WAI). The WAI [24] is a widely used and thoroughly studied instrument to measure work ability. The concepts of work ability, functional disability and illness behavior seems to be closely related [25]. The WAI was administered in sample E (all subjects).

Criterion Validity

Criterion validity is the extent to which a measure is related to an outcome. We examined the criterion validity of the WBI with regard to the following outcomes: clinical diagnosis, degree of disability associated with the sample (normal, sick leave, rehab), and prevalence of stressors associated with the sample (normal versus sick leave).

Clinical Diagnosis

Samples C and D were merged resulting in four diagnostics groups: anxiety disorder (n = 26), depressive disorder (n = 29), distressed/burnout (n = 22), and somatoform (chronic fatigue syndrome) (n = 21). First, we calculated the mean score to assess whether each diagnostic group obtained the highest score on the corresponding scale (e.g., depressed patients obtaining the highest score on the WBI depression scale). Using the normal sample as reference, T-scores ([z-score × 10] + 50) were calculated. T-scores ≥65 are commonly used as clinical cut-offs, and we hypothesized that the mean T-score score of the corresponding symptom scale for each clinical group should be 65 or higher. Next, a MANOVA was conducted with the diagnosis as the independent variable, and the four WBI symptom scales as the dependent variables to test the differences between the mean scale scores. Finally, to examine the diagnostic accuracy of the symptom scales, we calculated ROC-curves for the WBI-symptom scales. The diagnostic accuracy was examined with (a) regard to the ability of the WBI to detect a clinical diagnosis *within the normal population* and (b) with regard to the ability of the WBI to detect the proper diagnosis *within a clinical sample*. The stress scale was thus used to predict a distress/burnout diagnosis, the fatigue scale was used to predict the diagnosis a 'chronic fatigue syndrome', the anxiety scale was used to predict a diagnosis of 'anxiety disorder', and the depression scale was used to predict a diagnosis of 'depressive disorder.' Samples A, C and D were merged to examine (a) and sample C and D were used to examine (b). Differences in the area under the ROC-curve (AUC's) were tested using the procedure described by Hanley and McNiel [26].

Degree of Disability and Setting

We expected subjective disability and illness behaviors to vary with the type of setting. We hypothesized the lowest mean score for the WBI disability scale in the normal population (working employees, sample A), a moderate elevated mean score in the sick-listed employees (employees on sick leave sample, sample H), and a definite elevated score in rehab patients (sample G). The groups' differences in the mean score of the disability scale were tested with a one-way ANOVA test.

Prevalence of Stressors and Sick Leave Status

Since adversity in life and job strain could indirectly contribute to sick leave, we expected somewhat higher scores on the WBI life stressors and job strain scales in sick-listed employees compared to working employees. Substantial life stress and substantial job strain were defined as a score on the corresponding WBI scales of 1.5 SD (T ≥ 65) or higher. The hypothesis that the percentage of subjects with substantial life or job strain is higher among the sick-listed employees compared to working employees was tested using a χ^2-test.

Predictive Validity

Predictive validity is the extent to which a test predicts future outcomes. We tested the predictive validity of the WBI among a range of settings with varying degrees of disability: working employees, occupational health setting, and rehabilitation.

(a) In the normal population (sample B) we tested the ability of the WBI to predict future sick leave during (yes/no) during a period of 6 months after measurement. None of the subjects of sample B were on sick leave at the time of the measurement. Future sick leave was defined as: (1) 8 days or more of absenteeism irrespective the cause of the absenteeism; (2) more than six consecutive weeks of absenteeism because of a com-

mon mental disorder (CMD). The presence of a CMD was established by an occupational physician within 6 weeks.

(b) In the occupational health setting (sample H) we tested the ability of the WBI to predict medium (>1 month) and long-term absenteeism (>3 months) in patients who were on sick leave for 2 weeks at the time of measurement because of CMD or musculoskeletal symptoms without a specific medical cause.

(c) In the rehab setting (samples F, G) the WBI's predictive validity was examined by relating the pre-program WBI-scores to the work status (yes/no complete work resumption) at 3 months follow-up after treatment termination. Complete work resumption was defined as >90% working according to the pre-sickness contract hours.

First, all scale scores were transformed to T-scores (standardized scores) to facilitate a direct comparison between scales. Second, the scales were entered in a logistic regression analysis to discern the WBI scales' ability to predict the outcomes. Next, logistic regression analyses with stepwise backward selection (Wald) were conducted. With the stepwise procedure it is possible to eliminate scales sharing too much variance with the strongest predictor(s). Odds ratios were established to judge the predictive validity.

Results

Reliability

Internal Consistency and Test–Retest Reliability

Table 2 shows the internal consistency (Cronbach's alpha) and stability (test–retest reliability) of the 13 WBI scales. All scales exceed the minimal value of .70 for adequate internal consistency and test–retest reliability and most alpha's and r's are above .80.

Validity

Construct Validity

Item-rest correlations were performed as a first step in discerning the construct validity. Correlations (Pearson r) were calculated between each item and the sum of the other items of the scale (item-rest) and between the items and all other sales. In the normal sample three items failed (correlations with another scale is higher than the item-rest correlation). 'When I have finished work I feel really exhausted' (Job strain) correlated .01 higher with both the disability sale and the fatigue scale; 'I feel burnt out because of my work' (Job strain) correlated .05 higher with the disability scale; and 'Feeling 'rushed off your feet'' (Stress) correlated .02 higher with anxiety. However, in the patient sample the two job strain items did not fail. The stress item that failed in the normal sample also failed in the patient sample. Item 'Feeling 'rushed off your feet'' correlated .03 higher with anxiety and depression. Although at face value this item is really about stress, the data turned out otherwise.

Table 2 Internal consistency (Cronbach's alpha) and test–retest reliability (Pearson r) of the WBI-scales

WBI scale	Alpha (employees)[a]	Alpha (patients)[b]	Test–retest reliability (r)[c]
Social support at home	.78	.82	.78
Social support at work	.83	.86	.84
Job satisfaction	.84	.84	.84
Control	.84	.85	.89
Life stressors	.80	.79	.90
Job strain	.84	.86	.88
Avoidance	.86	.87	.82
Perfectionism	.79	.80	.83
Stress	.81	.82	.86
Fatigue	.84	.83	.86
Anxiety	.81	.83	.82
Depression	.81	.88	.76
Illness behavior	.89	.88	.84

[a]Combined samples A and B (n = 1.087)

[b]Combined samples C and D (n = 167)

[c]Subsample of sample A (n = 88)

The results of the confirmative factor analysis are presented in Table 3. In accordance with the structure of the WBI only the disability domain display an adequate fit for a 1-factor model. All other domains are not sufficiently explained by a 1-factor model, thus supporting the presence of underlying scales. The support and symptoms domains showed an adequate fit by a 4-factor model. The fit improved only slightly for the employees by the 4-factor bi-factor model, and could not be executed for the patients, which implies that the variance of the domain is largely explained by the 4 scales. The stressors and coping domains failed to be explained by a 2-factor model, whereas a 2-factor bi-factor model showed a moderate fit (CFI > .90). This implies that these two domains can be best understood as a combination of both underlying constructs (the scales) as well as a single construct (the domain itself). The pattern of the results is about the same for employees and patients thus supporting the validity of the use of the WBI in both normal subjects and patients.

Table 4 shows the unique variance of each WBI scale. Most scales demonstrate a substantial proportion of unique

Table 3 Results of confirmative factor analysis (CFA)

WBI domain	Employees			Patients		
	CFI[a]	TLI[a]	RMSEA[a]	CFI[a]	TLI[a]	RMSEA[a]
Support						
1-Factor model	.74	.72	.18	.70	.65	.20
4-Factor model	.92	.91	.10	.95	.94	.08
4-Factor bi-factor model	.95	.94	.08	na	na	na
Stressors						
1-Factor model	.70	.66	.18	.57	.50	.21
2-Factor model	.86	.83	.12	.89	.87	.10
2-Factor bi-factor model	.96	.94	.07	.95	.94	.07
Coping						
1-Factor model	.54	.47	.20	.55	.49	.22
2-Factor model	.83	.80	.12	.85	.83	.13
2-Factor bi-factor model	.92	.89	.09	.94	.92	.09
Symptoms						
1-Factor model	.91	.90	.12	.81	.79	.17
4-Factor model	.96	.95	.10	.95	.94	.09
4-Factor bi-factor model	.98	.97	.07	.98	.97	.08
Disability						
1-Factor model	.99	.99	.07	.97	.96	.16

CFI comparative fit index; *TLI* Tucker-Lewis index; *RMSEA* root mean square error of approximation

[a]CFI >.95, TLI >.95 and RMSEA <.06 are recommended as good models [10]

variance demonstrating the WBI scales' ability to measure distinct constructs. The proportion of unique variance of the symptoms scales and the illness behavior scale were lower in comparison with the other scales in the employees sample.

Table 5 shows the correlations (zero-order) and partial correlations between the five WBI domains. The pattern of partial correlations confirmed the expected weak association between the support domain and the other domains. The moderate size of the associations between the other domains corresponds with the assumed meaningful relationships between stressors, coping, and symptoms. The large zero-order correlations between the disability domain and the other domains could be largely explained by the level of symptoms, since the sizes of the partial correlations were substantially lower. Fisher's Z showed that the difference in absolute size of the partial correlations of the employees and patients samples were non-significant ($p > .05$) in 19 out of 20 pairwise comparisons. Thus, the pattern of correlations between the five WBI-domains is largely the same for employees and patients. Only the two sample's difference in partial correlations of the symptoms and disability domains was statistically significant (Fisher's $Z = 3.60$; $p < .001$).

Concurrent Validity

Table 6 presents the correlations between the WBI scales and related questionnaires. The highest correlation between each external measure and the concerning WBI scale is shown bold. Each WBI-scale showed the highest correlation with the most closely related construct. The highest correlation (.93) was between the WBI illness behavior scale and the Work Ability Index (WAI) which indicates that these scales measure almost the same. The WBI anxiety and depression scales showed substantial correlations (about

Table 4 Proportion of shared variance (R^2) and unique variance of the WBI-scales in normal subjects (employees)

WBI scales	Employees		Patients	
	R^2	Unique variance[a]	R^2	Unique variance[a]
Social support at home	.25	.53	.35	.47
Social support at work	.35	.48	.35	.51
Job satisfaction	.42	.42	.36	.48
Control	.22	.63	.13	.72
Life stressors	.27	.52	.33	.46
Job strain	.41	.43	.33	.53
Avoidance	.35	.51	.19	.68
Perfectionism	.19	.60	.22	.57
Stress	.60	.21	.50	.32
Fatigue	.57	.27	.35	.48
Anxiety	.52	.29	.31	.57
Depression	.54	.27	.52	.36
Illness behavior	.59	.29	.28	.60

[a]Unique variance = Cronbach's alpha (Table 3) − R^2

Table 5 Zero-correlations and partial correlations[a] between the WBI domains[b] of employees and patients

WBI domain	Support	Stressors	Coping	Symptoms	Disability
Support		−.31 (−.08)	−.15 (.07)	−.40 (−.20)	−.35 (−.08)
Stressors	−.25 (−.16)		.51 (.30)	.65 (.31)	.55 (.17)
Coping	.01 (.14)	.37 (.24)		.48 (.24)	.34 (−.05)
Symptoms	−.23 (−.15)	.55 (.36)	.40 (.25)		.72 (.54)
Disability	−.15 (−.03)	.34 (.15)	.13 (−.07)	.42 (.29)	

Correlations upper right side: employees (n = 912); Correlations bottom left side: patients (n = 167)

[a]Partial correlations between brackets. Each correlation between two domains is controlled by the other three domains

[b]Sum of the standardized scores of the underlying scales

.80) as well, indicating that they measure largely the same as the corresponding 4DSQ scales.

Criterion Validity

Clinical Diagnosis Table 7 shows the mean WBI symptom scales T-scores for four diagnostic groups of patients (distressed/burn-out, somatoform/chronic fatigue, anxiety disorder and depressive disorder). The highest mean T-score of each diagnostic group is shown bold. The hypothesis that each diagnostic group has the highest score mean T-score with its corresponding scale was confirmed. ROC-analyses revealed that all four scales successfully detected the four diagnoses among the general population of employees. That is, the scales discriminated perfectly patients with common psychopathology from working employees, and could also detect the proper diagnosis. The depression scale detected a depressive disorder adequately (AUC = .93), and the anxiety scale detected patients with an anxiety disorder fairly (AUC = .75). The stress and fatigue scales failed to detect the associated diagnosis (AUCs < .70). Thus, the stress and fatigue scales did not obtain enough unique variance to dis-

Table 6 Concurrent validity of the WBI scales; correlations (Pearson r) between the WBI scales and the other questionnaires

Questionnaires	Sample	Work and wellbeing inventory (WBI) scales												
		SSH	SSW	SAT	CTL	LST	JOB	AVO	PER	STR	FAT	ANX	DEP	ILL
Support														
JCQ–co-Worker support	A	.05	**.32***	.06	−.27	−.08	−.02	−.18	−.09	−.23	−.07	.13	−.22	−.25
JCQ–Supervisor support	A	.13	**.45***	.02	.13	.02	−.32*	.00	−.18	−.32*	.01	.07	−.09	−.37*
JCQ–Decision authority	A	−.01	.36*	.28	**.64***	.02	.10	−.27	.07	−.08	−.20	−.12	−.30	−.21
Stressors														
SRRS (life stressors)	A	−.32*	−.25	−.26	.10	**.46***	−.10	.23	.25	.45*	.40*	.43*	.40*	.40*
JCQ–Psychol. demands	A	.01	−.06	−.31*	.15	.37*	**.71***	.16	.62*	.31	.39*	.11	.09	.32*
Coping														
UCL–Avoidance	A	−.05	.04	−.11	−.02	.15	−.27	**.35***	−.04	.07	−.04	.11	.17	.07
UCL–Active tackling	A	.02	.04	.03	−.07	−.06	−.09	**−.32***	−.10	−.27	−.17	−.25	−.23	−.15
Symptoms														
4DSQ–Distress	D	−.33*	−.14	−.01	.18	.49	.13	.44	−.04	.71*	.77*	.56*	**.79***	.55*
4DSQ–Anxiety	D	−.15	.08	.01	−.02	.30	.05	.36	.04	.55*	.44*	**.81***	.65*	.40
4DSQ–Depression	D	−.38	−.20	−.21	−.01	.42*	.22	.40*	−.17	.50*	.55*	.52*	**.78***	.54*
4DSQ–Somatization	D	−.12	−.17	−.15	−.24	.15	.18	.29	−.11	.61*	**.66***	.60*	.55*	.65*
CIS20R (fatigue)	D	−.13	.11	.06	−.02	.19	.10	.28	−.08	.50*	**.57***	.33*	.47*	.53*
NRW (work-related fatigue)	D	−.44*	−.26	−.20	−.12	.16	.48*	.29	.18	.45*	**.53***	.14	.29	.32
Disability														
WAI (Work ability)	E	.16	.40*	.25*	.43*	−.51*	−.55*	−.56*	−.35*	−.79*	−.87*	−.79*	−.76*	**−.93***
FABQ–work	F	.14	.13	.18	.08	.16	.29	.20	.30	.14	.17	.28	.29	**.61***

Bold values represent the highest correlation between each external measure and the concerning WBI scale

SSH social support at home; *SSW* social support at work; *SAT* job satisfaction; *CTL* control; *LST* life stressors; *JOB* job strain; *AVO* avoidance; *PER* perfectionism; *STR* stress; *FAT* fatigue; *ANX* anxiety; *DEP* depression; *ILL* illness behavior

*p < .01

Table 7 Mean WBI symptom scales T-scores of patients[a] with common mental disorder (CMD) and area under curve (AUC)

WBI scale	Distressed/burn-out (n = 22) M (SD)	Somatoform (fatigue) (n = 21) M (SD)	Anxiety disorder (n = 26) M (SD)	Depression (n = 29) M (SD)	AUC[b] (95% CI)	AUC[c] (95% CI)
Stress	**68.6 (8.3)**	56.2 (10.4)	62.2 (7.5)	67.2 (9.0)	.91 (.86–.95)**	.68 (.55–.80)*
Fatigue	65.5 (9.8)	**70.0 (8.0)**	65.0 (10.0)	68.6 (8.8)	.91 (.88–.95)**	.60 (.46–.73)
Anxiety	59.8 (12.8)	51.0 (6.3)	**68.2 (6.2)**	63.7 (8.8)	.92 (.90–.95)**	.75 (.66–.84)**
Depression	61.5 (9.2)	52.0 (11.6)	59.8 (9.7)	**78.6 (8.4)**	.97 (.95–.98)**	.93 (.88–.98)**

Bold values represent the highest mean T-score of each diagnostic group

[a]Sample C (n = 98), [b]Normal subjects (sample A, n = 912) and patients (n = 98) together, [c]Only patients (n = 98)

*p < .05, **p < .001

criminate burn-out or fatigued patients from depressive or anxiety patients, whereas the other way around—with the depression and anxiety scales—depressive and anxious patients could be discriminated from burn-out and fatigued patients.

Degree of Disability and Setting The mean T-score of the WBI illness behavior scale was 50.2 (SD = 8.6) among the working employees (n = 912); T = 63.0 (SD = 16.2) in the group of sick-listed employees (n = 52), and T = 80.6 (SD = 13.6) among chronic back pain patients (n = 353) attending the rehab program. A performed one-way ANOVA showed the disability mean T-scores of the three groups to differ significantly (F = 1.057, p < .001).

Prevalence of Stressors and Sick Leave Status The prevalence of substantial life adversity as defined by a T-score ≥65 on the WBI life stressors scale was 7.7% among working employees and 34.6% among sick-listed employees ($\chi^2 = 42.6$; p < .001). The difference in prevalence of substantial job strain (T-score ≥65) between working employees (9.7%) and sick-listed employees (23.1%) was significant as well ($\chi^2 = 9.5$; p < .01). Hence, the hypothesis that life adversity and substantial job strain was more prevalent among sick-listed employees could be affirmed.

Predictive Validity

Table 8 shows the results of the logistic regression analyses in three different occasions of future risk.

Thirty-four employees (89%) were absent during 8 days or more during a period of 6 months. Fatigue turned out to be the strongest predictor of future absenteeism in working employees after the stepwise backward procedure. Interestingly, other variables predicted which subjects would be off work during for 6 weeks or more because of a common mental disorder (CMD). This happened only in seven subjects (2%). The level of depression was clearly the strongest predictor. Symptoms and disability predicted which 2-weeks sick listed employees because of CMD or musculoskeletal symptoms would stay on sick leave for a longer period (>1 and >3 months). The stepwise procedure revealed a substantial shared variance between symptoms and disability. The level of illness behavior turned out to be the most robust predictor. Prediction of work resumption in rehab patients showed modest results. After the stepwise procedure, only illness behavior was left as a significant predictor for work resumption in rehab patients.

Discussion

Main Findings

The results of this study show that the psychometric properties of the WBI are good. All scales turned out to be reliable (test–retest reliability and internal consistency). Several features of validity were examined: construct, concurrent, criterion and predictive validity. The construct validity of the WBI could be clearly demonstrated. The structure of the WBI was confirmed by confirmative factor analysis, and the domains of the WBI showed the expected and meaningful relationships. Furthermore, the 13 scales of the WBI overall showed high unique variances, which means that each scale taps at least some unique construct. The lowest unique variance was found for the stress scale, which was not a surprise as stress is the most common response to any adversity, and will therefore easily overlap with the other scales. The WBI showed good concurrent validity for most of its scales. The high correlation between the illness behavior scale and the WAI illustrates that these two scales measure about the same construct. Regarding the criterion related validity, we found that the illness behavior scale could make a distinction between samples with different levels of disability (healthy employees, sick employees and rehab patients). The symptom scales fairly distinguished depressed patients from distressed patients. The predictive validity of the WBI concerning future sickness absence and return to work could be

Table 8 Logistic regression analyses with future sick-leave, long term absenteeism and work resumption as dependent variables and the WBI scales (T-scores) as predictor variables

WBI scales	Future sick leave in non-absent working employees (n = 388) during a period of 6 months		Long term absenteeism in employees (n = 132) at 2-week sick leave because of CMD or musculoskeletal disorder		Work resumption of rehab patients (n = 773) chronic back pain (n = 323) fatigue (n = 450)	
	>1-week sick leave	>6 weeks sick leave because of CMD	1-month sick leave	>3-month sick leave	Work resumption (<50%) at 3-month follow-up	Work resumption (<50%) at 3-month follow-up
	(n = 34; 9%)	(n = 7; 2%)	(n = 96; 73%)	(n = 51; 39%)	(n = 60; 19%)	(n = 60; 13%)
	Odds ratio[a] (95% CI)	Odds ratio[a] (95% CI)	Odds ratio[a] (95% CI)	Odds ratio[a] (95% CI)	Odds ratio[a] (95% CI)	Odds ratio[a] (95% CI)
Independent predictors						
Support						
Social support at home	0.91–1.00	0.84–1.02	0.93–1.03	0.92–1.03	0.95–1.01	0.99–1.03
Social support at work	0.93–1.02	0.89–1.07	0.94–1.03	0.96–1.05	0.95–1.01	0.96–1.02
Job satisfaction	0.96–1.06	0.85–1.00	0.95–1.04	0.96–1.05	0.95–1.02	0.94–1.02
Control	0.91–1.00	0.83–1.01	0.94–1.03	0.94–1.04	0.96–1.02	0.96–1.02
Stressors						
Life stressors	0.97–1.07	0.90–1.12	0.98–1.07	0.98–1.08	0.98–1.03	0.99–1.04
Job strain	0.95–1.05	0.93–1.14	0.99–1.09	0.99–1.09	0.97–1.06	0.93–1.00
Coping						
Avoidance	1.00–1.08	1.04–1.19	0.97–1.07	0.97–1.06	0.97–1.01	0.98–1.02
Perfectionism	0.99–1.07	0.98–1.16	0.98–1.06	0.98–1.06	0.97–1.02	0.97–1.02
Symptoms						
Stress	1.01–1.10	1.04–1.23	1.01–1.11	0.98–1.07	0.99–1.03	0.98–1.03
Fatigue	1.01–1.09	1.04–1.20	1.02–1.13	1.00–1.09	1.00–1.05	1.01–1.07
Anxiety	0.99–1.09	1.04–1.24	1.01–1.10	0.99–1.08	0.98–1.02	0.98–1.02
Depression	0.98–1.12	1.07–1.31	1.00–1.10	0.99–1.08	1.00–1.04	0.97–1.01
Disability						
Illness behavior	1.01–1.08	1.00–1.15	1.04–1.13	1.01–1.10	1.03–1.09	1.03–1.08
Final model						
Job strain	–	–	1.01–1.10	–	–	–
Fatigue	1.01–1.09	–	–	–	–	–
Depression	–	1.07–1.31	–	–	–	–
Illness behavior	–	–	1.01–1.11	1.01–1.10	1.03–1.09	1.03–1.08

CMD common mental disorders

[a]The odds ratio is related to the WBI scale's T-score

established for future sickness absence in healthy workers, return to work of sick-listed workers, and return to work after rehabilitation. The WBI seems to be most successful in predicting return to work in sick-listed workers.

Comparison with Other Studies

In the present study, particularly the illness behavior scale appeared to be the most robust predictor for the various work related outcomes, and particularly for return-to-work self-efficacy. This finding is in line with the research literature that reports that people experiencing health complaints that truly hinder their functioning in life, will not recover quickly or return back to work [27, 28].

Overall, the literature about relevant predictors of future sickness absence varies to some extent [28–34]. For example, some studies report that experiencing fatigue is an important predictor for future sickness absence [29, 30]. However, fatigue was not found to be associated with sickness absence in women in two other studies [31, 32]. Whereas in the study by Norder et al. fatigue, measured as the need for recovery, was found to be a predictor for future

risk of developing depression [33]. In the study by van Hoffen et al. stress as measured with the 4DSQ distress scale was the strongest predictor for future long term sick leave due to mental health complaints [34]. In our study, fatigue could, to some extent, predict future sickness absence, but depression was the best predictor for future sickness absence because of mental health complaints. For a part these differences in predictors may be explained by the overlap between mental health symptoms. As was previously reported by Terluin et al., the 4DSQ distress scale shares 35–45% of its variance with the other 4DSQ scales, particularly with the depression scale [11]. Also, in our study, we found that stress symptoms shared most variance with the other symptoms scales, and as a result had the lowest unique variance.

Strengths and Limitations

A strong feature of this study is that the construct validity was tested in both patient as well as in working employees. Instruments are often constructed within one population (e.g., healthy or aging employees), and then applied or further validated in patient samples or vice versa. The validation of the WBI in both healthy workers and patients guarantee a more solid ecological validity of the test.

We also want to address some limitations of the present study. In both samples the stress item 'feeling rushed of your feet' correlated more with the constructs of anxiety and depression instead of the construct of stress. Although, at face value this item is really about stress, apparently this item combines the emotional turmoil someone experiences when there are feelings of anxiety. This item therefore needs to be reconsidered in future versions of the WBI. Also, two items of the job strain scale in the normal sample failed in this respect. Apparently, healthy employees who experience feelings of exhaustion after work, or of burn-out due to work, relate these feelings more to fatigue and disability in general, than to experienced job strain. However, for patients these items of job strain did correlate well with the construct of job strain. It remains debatable whether these items, within the construct of job strain, need to be adjusted risking loss of reach in measuring job strain for patients. We decided to keep this item in order to establish a job strain scales with a broad range that is applicable in both patients and working employees without any floor or ceiling effects.

Another limitation of this study is the limited sample size (n = 388) of healthy employees. Although the percentage of subjects (1.8%) developing a new common mental disorder within 6 months seems consistent with literature, in absolute terms (seven subjects) the amount of subjects is too low to build a prediction model [34]. In that sense, these results must be regarded as very thin evidence in need for further validation using larger samples sizes.

The prediction of return to work in rehabilitation patients seems at face value disappointing. We might conclude that the WBI is of little use in the assessment of chronic pain patients in a rehab setting. This conclusion might be premature however. One must realize that the rehab program was largely focused on addressing and treating maladaptive illness behavior and other psychosocial problems. If a construct is the central focus of treatment than the predictive power of this same construct used as a predictor of treatment outcome will of course diminish. Whether this line of reasoning is true needs further study.

Finally, the validity and reliability of the WBI have been tested over the years in the Dutch setting with Dutch patients and employees. The generalizability of the WBI towards other national contexts and cultures may therefore be limited. On the other hand, based on the existing literature, the concepts used in the WBI are already frequently researched in other countries. To what extent the WBI is also valuable for workers without a particular employment contract such as independent workers needs to be established in future studies.

Practical Implications and Further Research

One of the most attractive parts of the WBI for occupational health practice is the prediction of a future course for different work related outcomes. Although, some other questionnaires and screening tools already exist to predict future sickness absence or the risk of long term sickness absence, these predicting tools mainly exist for particular diseases, within particular working populations or for single work related outcomes [1, 35]. The WBI has the potential to predict different outcome variables within different type of patients and for healthy employees together with, and an assessment of, relevant work and home life related topics by means of scales.

The unique variance of symptom scales such as stress, fatigue, anxiety and depression was higher for patients than for healthy employees. One explanation for this observation could be that with the progress of time the relationship between stressors and symptoms becomes less tight. In working employees, a slight increase of psychosocial problems may result in a proportional increase in symptoms, whereas in patients there seems to be a less direct relationships between problems (e.g., job strain, life stressors) and symptoms. As stated earlier, with ongoing time, symptoms start their 'own dynamics'. For example, some patients with only minor psychosocial problems will start worrying about the symptoms which may result in an even further increase of symptoms. Such processes may blurry the 'proportional' relationship between problems and symptoms in patients. Also medical conditions may affect the mental symptoms in patients which is not the case in working employees. For clinical practice the implication of this could be that

at different moments in the process of absenteeism other predictive variables are relevant.

The most important value of the WBI is that it provides the occupational physician a broad screening of relevant topics, topics that cover the whole range from work- and life-related issues that facilitates to make a quick scan about which issues are relevant and which not. More research is needed to further validate the WBI. Especially larger samples sizes are needed to provide a more solid base of the results provided in this study.

Conclusion

The WBI is a valid and reliable tool for occupational health practitioners to screen for risk factors for prolonged or future sickness absence. With this tool they will have indications for further advice and interventions.

Compliance with Ethical Standards

Conflict of interest A. A. Vendrig has developed the WBI (Dutch: VAR-2) and he participates in a project to support the online application of the VAR-2 (project: VAR-2-app). F. G. Schaafsma declares she has no conflicts of interests.

Research Involving with Human and Animal Participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

Appendix

See Table 9

Table 9 Summary of the WBI scales and subscales

Scale (number of items)	Scale description (short description, subscales, examples of items, main theory)
Social support at home (5)	*Short description*: assesses the level of perceived social support in one's home environment. *Examples of items*: 'my family or friends give me good advice' and 'there are people in whom I can confide and they listen.' *Main theory*: buffer hypothesis of social support [36]
Social support at work (5)	*Short description:* assesses the level of perceived social support at work. *Subscales*: 'social support manager or supervisor' and 'social support colleagues'. *Examples of items:* 'I get along well with my manager or supervisor' and 'I feel valued by my colleagues.' *Main theory*: job-demand-control-support model [37]
Job satisfaction (7)	*Short description:* this scale measures how satisfied one is with one's current job. *Subscales*: 'work experience', 'challenge', 'competence', 'terms and condition of employment' and 'considering changing jobs.' The subscales helps to pinpoint the exact reasons of low job satisfaction if applicable. *Examples of items:* 'I thoroughly enjoy my work' and 'I am thinking of finding another job.' *Main theory*: job characteristics theory [38]
Control (5)	*Short description:* this scale measures how much control and freedom the person experiences in his or her work. *Examples of items*: 'I can decide when to take my breaks at work' and 'I can plan my working day the way I want to.' *Main theory*: job-demand-control model [39]
Life stressors (8)	*Short description:* the presence of problems in one's life is assessed by this scale. *Subscales*: 'life events' and 'daily hassles.' *Examples of items:* 'major changes have occurred in my private life in the past year' and 'my home life is filled with arguing and bickering.' *Main theory*: social origins of depression [40]
Job strain (8)	*Short description:* she scale assesses the level of experienced job strain both in terms of traditional work load and in terms of 'modern' utterances of job strain like emotional burden. Indicators of job strain like the need for recovery and being unable to switch off are also included. *Subscales:* 'workload', 'emotional burden', 'need for recovery', and 'switch off.' *Examples of items:* 'I need to work very hard in order to finish my daily tasks' and 'when I have finished work I feel really exhausted.' *Main theory*: job-demand-control model [39]
Avoidance (8)	*Short description:* persons obtaining high scores are insecure, avoids conflicts and postpone difficulties. *Subscales:* 'avoidance behavior' and 'problem solving skills.' *Examples of items:* 'I wait a long time before I express my irritations' and 'it is difficult for me to ask for help if I need it'. *Main theory*: BIG-5 model of personality [41]
Perfectionism (9)	*Short description:* persons obtaining high scores put extreme demands on themselves and others and often cross their own limits. *Subscales:* 'insistent' and 'diligence.' *Examples of items:* 'people who cut corners irritate me a lot' and 'other people believe I am a perfectionist.' *Main theory*: BIG-5 model of personality [41]
Stress (6)	*Short description:* this scale assesses the most common stress symptoms. *Examples of items:* 'become easily annoyed or irritated' and 'feeling 'rushed off your feet''. *Main theory*: general adaption syndrome (GAS) [42]
Fatigue (4)	*Short description:* the fatigue scale assesses abnormal/clinical fatigue (in contrast to normal fatigue) like the absence of normal recovery after minimal effort and being totally exhausted. In normal samples fatigue is hard to separate from stress. In clinical samples, however, fatigue seems to be a separate factor. In the context of work, fatigue is a much common reason for absenteeism. *Examples of items:* 'not feeling fit and rested in the morning' and 'feeling tired after minimal activity.' *Main theory*: fatigue at work [29]

Table 9 (continued)

Scale (number of items)	Scale description (short description, subscales, examples of items, main theory)
Anxiety (5)	*Short description:* includes clinical manifestations of anxiety states. *Examples of items:* 'feeling nervous' and 'feeling worries or uneasy about things that could happen.' *Main theory*: tripartite model of anxiety and depression [43]
Depression (6)	*Short description:* contains clinical symptoms of depressions. *Examples of items:* 'having no interest in anything' and 'feeling that life is meaningless.' *Main theory*: tripartite model of anxiety and depression [43]
Illness behavior (7)	*Short description:* this scale assess the extent to which the symptoms (those symptoms the person are bothering the most, irrespective whether they are psychic or physical) are interfering with one's daily functioning. *Subscales:* 'impact of symptoms', 'need to rest', 'own prognosis of recovery', 'symptom aggravation due to work', 'threshold for work resumption.' *Example of items:* 'it will take a long time before I am fully recovered' and 'it is not sensible to continue to work with my symptoms.' *Theory*: fear-avoidance beliefs model [44]

References

1. Steenstra IA, Munhall C, Irvin E, Oranye N, Passmore S, Van Eerd, Mahood Q, Hogg-Johnson S. Systematic review of prognostic factors for return to work in workers with sub acute and chronic low back pain. J Occup Rehabil. 2016. doi:10.1007/s10926-016-9666-x.
2. Nieuwenhuijsen K, Verbeek JH, Boer AG de, et al. Predicting the duration of sickness absence for patients with common mental disorders in occupational health care. Scand J Work Environm Health. 2006;32(1):67–74.
3. Corbière M, Negrini A, Durand MJ, St-Arnaud L, Briand C, Fassier JB, Loisel P, Lachance JP. Development of the return-to-work obstacles and self-efficacy scale (roses) and validation with workers suffering from a common mental disorder or musculoskeletal disorder. J Occup Rehabil. 2016. doi:10.1007/s10926-016-9661-2.
4. Kristman VL, Shaw WS, Boot CRL, Delclos GL, Sullivan MJ, Ehrhart MG. Researching complex and multi-level workplace factors affecting disability and prolonged sickness absence. J Occup Rehabil. 2016;26(4):399–416.
5. Bakker AB, Demerouti E. The job demands–resources model: state of the art. J Man Psychol. 2007;22(1):309–328.
6. Klink JJL, Ausems CMM, Beijderwellen BD, et al. Evidence-based Occupational Health Guideline on common mental health for workers (Richtlijn handelen van de bedrijfsarts bij werkenden met psychische problemen). Utrecht: NVAB; 2007.
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington: American Psychiatric Publishing; 2013.
8. Verschuren CM, Nauta AP, Bastiaanssen MHH. Multidisciplinary guideline on mental stress and burnout for primary and occupational care professionals (Multidisciplinaire richtlijn overspanning en burnout voor eerstelijns professionals). Utrecht: NVAB/NHG/LVE; 2011.
9. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess. 2003;80(1):99–103.
10. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equat Model. 1999;6(1):1–55.
11. Terluin B, van Marwijk HWJ, Adèr HJ, de Vet HCW, Penninx B, Hermens M, van Boeijen CA, Balkom A, van der Klink JJL, Stalman WAB. The four-dimensional symptom questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. BMC Psychiatry. 2006;6(1):1–20.
12. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res. 1994;38(5):383–392.
13. Beurskens AJ, Bültmann U, Kant I, Vercoulen JH, Bleijenberg G, Swaen GM. Fatigue among working people: validity of a questionnaire measure. Occup Environm Med. 2000;57(5):353–357.
14. Vendrig AA, Deutz P, Vink I. Dutch translation and addaptation of the fear-avoidance beliefs questionnaire (Nederlandse vertaling en bewerking van de fear-avoidance beliefs questionnaire. Ned Tijd Pijn Pijnbestrijd. 1998;18(1):11–14.
15. Waddell G, Newton M, Henderson I, Somerville D, Main C. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance in chronic low back pain and disability. Pain. 1993;52(2):157–168.
16. Houtman I. Reliability and validity of the Dutch version of the karasek job content questionnaire. In: NIOSH/APA conference on stress, work and health. Washington, DC: APA; 1995.
17. van der Doef M, Maes S. The job demand–control (-support) model and psychological well-being: a review of 20 years of empirical research. Work Stress. 1999;13(2):87–114.
18. van Veldhoven M, Broersen S. Measurement quality and validity of the need for recovery scale. Occup Environm Med. 2003;60(Suppl 1):i3–i9.
19. Sluiter JK, de Croon EM, Meijman TF, Frings-Dresen MH. Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. Occup Environm Med. 2003;60(Suppl I):i62–i70.
20. Holmes TH, Rahe RH. The social readjustment scale. J Psychosom Res. 1967;11(2):213–218.
21. Gerst MS, Grant I, Yager J, Sweetwood H. The reliability of the social readjustment rating scale: moderate and long-term stability. J Psychos Res. 1978;22(6):519–523.
22. Schreurs PJG, van Willige G, Tellegen B, Brosschot JF. Revised instruction manual fort he Utrecht Coping List (Herziene handleiding Utrechtse Coping Lijst (UCL)). Lisse: Swets & Zeitlinger; 1993.
23. Turner H, Bryant-Waugh R, Peveler R, Bucks R. A psychometric evaluation of an English version of the Utrecht Coping List. Eur Eat Disord Rev. 2012;20(4):339–342.
24. Tuomi K, Ilmarinen J, Jahkola A, et al. Work ability index. 2nd revised ed. Helsinki: Finnish Institute of Occupational Health; 1998.
25. van den Berg TIJ, Elders LAM, de Zwart BCH, Burdorf A. The effects of work-related and individual factors on the Work Ability Index: a systematic review. Occup Environm Med. 2009;66(4):211–220.
26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 1983;148(3):839–843.
27. Volker D, Zijlstra-Vlasveld MC, Brouwers EP, van Lomwel AG, van der Feltz-Cornelis CM. Return-to-work self-efficacy and

actual return to work among long-term sick-listed employees. J Occup Rehabil. 2015;25(2):423–431.
28. Momsen AH, Stapelfeldt CM, Nielsen CV, Nielsen MB, Rugulies R, Jensen C. Screening instruments for predicting return to work in long-term sickness absence. Occup Med. 2016;67(2):101–108.
29. Janssen N, Kant IJ, Swaen GW, Janssen PP, Schröer CA. Fatigue as a predictor of sickness absence: results from the Maastricht Cohort study on fatigue at work. Occup Environm Med. 2003;60(Suppl I):i71–i76.
30. Leone SS, Huibers MJ, Knottnerus JA, Kant I. The prognosis of burnout and prolonged fatigue in the working population: a comparison. J Occup Environm Med. 2008;50(10):1195–1202.
31. Bültmann U, Nielsen MB, Madsen IEH, Burr H, Ruglies R. Sleep disturbances and fatigue: independent predictors of sickness absence? A prospective study among 6538 employees. Eur J Public Health. 2013;23(1);123–128.
32. Roelen CA, van Rhenen W, Groothoff JW, van der Klink JJ, Bültmann U. Prolonged fatigue is associated with sickness absence in men but not in women: prospective study with 1 year follow-up of white collar employees. Int Arch Occup Environm Health. 2014;87(3):257–263.
33. Norder G, Roelen CA, van der Klink JJ, Bültmann U, Sluiter JK, Nieuwenhuijsen K. External validation and update of a prediction rule for the duration of sickness absence due to common mental disorders. J Occup Rehabil. 2016;27(2):202–209.
34. van Hoffen MFA, Joling CI, Heymans MW, Twisk JWR, Roelen CAM. Mental health symptoms identify workers at risk of long term sickness absence due to mental disorders: prospective cohort study with 2-year follow up. BMC Public Health 2015;15(1):1235.
35. Boezeman EJ, Nieuwenhuijsen K, Sluiter JK. Predictive value and construct validity of the work functioning screener-healthcare (WFS-H). J Occup Health. 2016;58(2):163–169.
36. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull. 1985;98(2):310–357.
37. Johnson JV, Hall EM. Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population. Am J Public Health. 1988;78(10):1336–1342.
38. Hackman JR, Oldham GR. Development of the job diagnostic survey. J Applied Psychol. 1975;60(2):159–170.
39. Karasek RA. Job demands, job decision latitude and mental strain, implications for job redesign. Adm Sci Q. 1979;24(2):285–308.
40. Brown GW, Harris TO. Social orgins of depression. London: Tavistock; 1978.
41. Costa PT, McCrae RR. Domains and facets: hierarchical personality assessment using the Revised NEO Personality Inventory. J Pers Assess. 1995;64(1):21–50.
42. Selye H. Stress and the general adaptation syndrome. Br Med J. 1950;1(4667):1383–1392.
43. Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. J Abnormal Psychol. 1991;100(3):316–336.
44. Vlaeyen JW, Linton SJ. Fear-avoidance beliefs and its consequences in chronic muscoloskeletal pain: a state of the art. Pain. 2000;85(3):317–332.