

Original

Predictive value and construct validity of the work functioning screener-healthcare (WFS-H)

Edwin J. Boezeman¹, Karen Nieuwenhuijsen¹ and Judith K. Sluiter¹

¹Coronel Institute of Occupational Health, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands

Abstract: Objectives: To test the predictive value and convergent construct validity of a 6-item work functioning screener (WFS-H). **Methods:** Healthcare workers (249 nurses) completed a questionnaire containing the work functioning screener (WFS-H) and a work functioning instrument (NWFQ) measuring the following: cognitive aspects of task execution and general incidents, avoidance behavior, conflicts and irritation with colleagues, impaired contact with patients and their family, and level of energy and motivation. Productivity and mental health were also measured. Negative and positive predictive values, AUC values, and sensitivity and specificity were calculated to examine the predictive value of the screener. Correlation analysis was used to examine the construct validity. **Results:** The screener had good predictive value, since the results showed that a negative screener score is a strong indicator of work functioning not hindered by mental health problems (negative predictive values: 94%-98%; positive predictive values: 21%-36%; AUC: .64-.82; sensitivity: 42%-76%; and specificity 85%-87%). The screener has good construct validity due to moderate, but significant ($p < .001$), associations with productivity ($r = .51$), mental health ($r = .48$), and distress ($r = .47$). **Conclusions:** The screener (WFS-H) had good predictive value and good construct validity. Its score offers occupational health professionals a helpful preliminary insight into the work functioning of healthcare workers.

(J Occup Health 2016; 58: 163-169)

doi: 10.1539/joh.15-0056-OA

Key words: Health, Work, Employee performance appraisal, Work functioning, Brief questionnaire, Cross-sectional survey

Introduction

Nurses, medical doctors, and other types of healthcare workers regularly have a high workload, low job autonomy, and frequent interactions with suffering patients¹. Many healthcare workers also have irregular working hours and are held highly accountable for errors made during work. Consequently, mental health problems such as stress, anxiety, and depression are highly prevalent among healthcare workers². Studies show that mental health problems undermine the ability of healthcare workers to contribute to the recovery process and quality of life of recipients of medical care services³. Thus, it is critical to regularly screen and evaluate the work functioning of healthcare workers. Work functioning refers to the health-related ability of workers to meet their work demands⁴. Researchers have developed screening instruments for detecting mental health issues such as depression⁵, distress^{6,7}, and posttraumatic stress disorder⁸. However, a brief screener instrument for evaluating the work functioning of healthcare workers is not yet available. Researchers previously used the 47-item Nurses Work Functioning Questionnaire (NWFQ) as a screening tool for evaluating the work functioning of healthcare workers⁹. However, it is inefficient to use long questionnaires for screening the work functioning of healthcare workers since relatively long screening procedures keep healthcare workers too long from their jobs. Furthermore, a relatively long screening procedure may yield incomplete information as questionnaire length is a determinant of unwillingness of individuals to complete the questionnaire¹⁰. Thus, there is a need for a brief instrument for screening and evaluating the work functioning of healthcare workers.

Received February 24, 2015; Accepted October 17, 2015

Published online in J-STAGE March 24, 2016

Correspondence to: K. Nieuwenhuijsen, (Ph.D.), Coronel Institute of Occupational Health, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands (e-mail: K.Nieuwenhuijsen@amc.nl)

Screening for unimpaired work functioning

Screeners are brief measurement instruments that offer an indication of the absence or presence of disease or impairment in individuals. These instruments are helpful for gaining a preliminary insight into the functioning of individuals. They also are useful for identifying individuals eligible for further examination due to a potential need for care or support. Items for screeners can either be derived from refined (longer) instruments that measure the concept that the screener should capture or they can be newly developed items. Researchers have addressed the predictive value of screeners containing single items^{5,8,11}) and screeners containing up to 6-10 items⁶). The effectiveness of screeners is primarily evaluated in terms of their predictive value and less in terms of their internal consistency and structural validity. Sensitivity, specificity, size and significance of area under the curve (AUC), and positive and negative predicted values are psychometric properties that indicate the predictive value or effectiveness of a screener measure⁵).

Aim of the research

The aim of the current research is to examine the predictive value and convergent construct validity of the newly developed 6-item Work Functioning Screener-Healthcare (WFS-H). It is reasoned that a negative WFS-H test score indicates sufficient ability to (continue to) fulfill the work demands that are difficult to complete for individuals with mental health problems. The WFS-H is job-specific and thus different from generic instruments, such as the Work Limitations Questionnaire (WLQ)¹²) and the Work Role Functioning Questionnaire (WRFQ)¹³). The WFS-H is specifically relevant for measuring the work functioning of healthcare workers working in the healthcare sector, since the WFS-H takes into consideration the work circumstances under which healthcare workers are expected to meet their work demands (e.g., work context that is unpredictable and hectic and work context in which errors have severe consequences). Furthermore, due to its short length, it is more efficient to use the WFS-H for screening the work functioning of healthcare workers than to use existing (longer) instruments, such as the NWFQ.

Methods

Data and procedure

The data of the current research (249 nurses; 40 males, 209 females; mean age=47, SD=12; mean hours employed=30, SD=6.3) were retrieved from the dataset collected by Gärtner et al.¹⁴) in the context of their study on the psychometric properties of the NWFQ. Their dataset contained a set of screener items that were developed independently from the NWFQ, yet the predictive value of those screener items was left unexamined. The research

had a cross-sectional survey design, and 1200 healthcare workers (i.e., 1050 nurses and 150 allied health professionals) were invited to participate in the research. The Medical Ethical Committee of the Academic Medical Center approved the research. The dataset contained 249 nurses (response rate: 24%) and 58 allied health professionals (response rate: 39%) who had completely filled in a questionnaire and had indicated their worker type. The group of allied health professionals was found to be too small for conducting subgroup analyses. At the same time, at this stage of the research, it was not an option to combine the two groups into one sample. Furthermore, for the current research, the NWFQ was selected as a “gold standard” instrument, yet Gärtner et al.⁹) concluded that not all subscales of the NWFQ are applicable to allied health professionals. For these reasons, only the data of the nurses (n=249) were used for the current research. See Gärtner et al.¹⁴) for further details on the data collection procedure, including informed consent and ethical approval. The demographic characteristics (age, number of hours employed, type of nursing job, etc.) of the research participants were determined through general questions.

Screener for detecting impaired work functioning

The 6-item work functioning screener (WFS-H) contained the following self-developed items: “My work functioning is worse than before,” “Without the help of my colleagues I would not function adequately,” “I make mistakes more often,” “I take risks more often in my job than before,” “I retire to a less hectic place more than before,” and “For a while now I have had less interest for my work activities.” Studies^{3,9,14}) clarify that the conceptual content of these items reflects how mental health problems typically affect healthcare workers and their work performance (e.g., mental health problems make it more difficult to concentrate and stay calm at work, motivate to take extended breaks from work or to report sick to momentarily leave the workplace, undermine work motivation, increase risk-taking since they for instance associate positively with careless handling of needles, etc.). The scale of the brief screener used a 4-point answer format (0=not at all, 1=slightly agree, 2=agree, 3=totally agree). For each item, the answer option “not at all” reflects unimpaired work functioning. For an indication of the work functioning of healthcare workers it is essential to sum the item scores. The raw screener sumscore ranges from 0 to 18. A score of 0-100 can be calculated using the following formula: $\text{Score}_{0-100} = ((\text{raw sumscore WFS-H} \cdot 100)/18)$. For an overview, the items of the WFS-H are included in the Appendix.

Measure to examine the predictive value of the screener

The NWFQ⁹) was used to test the predictive value of the brief work functioning screener (WFS-H). The

NWFQ is a questionnaire developed to measure in nurses and allied health professionals levels of impaired work functioning due to mental health problems. The aspects of work functioning that the NWFQ measures are as follows: cognitive aspects of task execution and general incidents (11 items), causing incidents at work (8 items), avoidance behavior (8 items), conflicts and irritations with colleagues (7 items), impaired contact with patients and their family (8 items), and lack of energy and motivation (5 items). The items of the NWFQ have a reference period of 4 weeks, and for each subscale the item scores are summed and transformed to a 0-100 scale with the following formula: $((\text{sum of items} \times 100) / (\text{number of subscale items} \times \text{maximum item score}))$. For each subscale, a high score is indicative of impaired work functioning due to mental health problems. Studies¹⁵⁾ show that the NWFQ has good psychometric properties. For instance, the internal consistency of the NWFQ subscales used in the current research is good ($\alpha > .70$), and these subscales also have good construct validity and good discriminative validity. Cut-off points exist in the literature for each subscale of the NWFQ^{16,17)}, and a score above the cut-off point indicates impaired work functioning due to mental health problems. The cut-off point of the subscale “cognitive aspects of task execution and general incidents” is 25, “avoidance behavior” is 13, “conflicts and irritations with colleagues” is 29, “impaired contact with patients and their family” is 19, “lack of energy and motivation” is 32, and “causing incidents at work” is 15.

Measures to examine the construct validity of the screener

Empirical support for predictions that the score of the measure under examination correlates strongly, but not too strongly, with scores of measures that also capture the construct of interest indicates good convergent construct validity¹⁸⁾. The WFS-H is an instrument for determining whether workers meet work demands that are difficult to meet for individuals with mental health problems. Studies show that mental health problems hinder worker productivity¹⁹⁾. Hence, the 25-item Endicott Work Productivity Scale (EWPS)²⁰⁾ was used for recording worker productivity. This instrument uses a 5-point Likert scale for recording productivity (1=never, 5=always), and a high score suggests productivity loss. The EWPS has good internal consistency ($\alpha > .70$) and good test-retest reliability (ICC=0.92). Further, the 12-item version of the General Health Questionnaire (GHQ)^{21,22)} and the distress scale (16 items) of the Four Dimensional Symptoms Questionnaire (4DSQ)^{23,24)} were also used to test for convergent construct validity. Many different types of mental health problems (depression, anxiety disorder, etc.) affect work functioning, and these instruments record symptoms that are characteristic for multiple types of mental health problems. The GHQ uses a 4-point Likert scale (1=not at all, 4=

more than usual); an overall mental health score is calculated by summing the item scores, and a high score is suggestive of mental health problems. The GHQ has, for instance, good internal consistency ($\alpha > .70$). The distress-scale of the 4DSQ has 16 items and uses a 5-point Likert scale for recording the responses to the items (1=no, 5=very often). A high summed score suggests high distress, and the scale has good internal consistency ($\alpha > .70$).

Analyses

First, the raw subscale scores of the NWFQ and the raw WFS-H score were transformed to 0-100 scores. Subsequently, the NWFQ subscale scores were categorized into positive test results (i.e., score above cut-off point) and negative test results (i.e., score below the cut-off point). Then, ROC analysis²⁵⁾ conducted in SPSS (IBM Corp., Armonk, NY, USA) was used to determine whether the WFS-H could serve as a test for the NWFQ-subscale (AUC > .50, $p < .05$). When this appeared the case, the ROC-analysis suggested the optimal cut-off point for the WFS-H at the point of optimal balance between sensitivity and specificity²⁵⁾. For the WFS-H, a score above the cut-off point represented a positive test result and a score below the cut-off point represented a negative test result. Cross-tabulations subsequently revealed frequencies of cases of test match and test mismatch between the WFS-H and the NWFQ, and all these frequencies were entered into MedCalc. MedCalc is a computer program that uses algorithms developed by Meraldo and Lau²⁶⁾ for calculating sensitivity, specificity, and negative and positive predictive values. After data entry, MedCalc yielded the sensitivity, specificity, and the negative and positive predictive values of the WFS-H. ROC analysis conducted in SPSS was subsequently used for calculating the final AUC values. The negative predictive value indicates the likelihood that a worker who has a negative WFS-H score has unimpaired work functioning. A significant AUC, i.e., exceeding .50, indicates that the WFS-H has value as a test. Correlation analysis was used in testing for construct validity. It was predicted that the WFS-H scores would correlate moderately ($r = .30-.60$) with respectively the scores of the EWPS, the GHQ, and the distress-subscale of the 4DSQ.

Results

The ROC analyses for determining the cut-off point of the WFS-H revealed in terms of significant AUCs ($> .50$; $p < .05$) that the WFS-H could serve as a test for five of the six work functioning domains that the NWFQ measures (Fig. 1). These work functioning domains were: cognitive aspects of task execution and general incidents (mean=9.3, SD=12.6; minimum observed value=0, maximum observed value=82); avoidance behavior (mean=4.7, SD=9.5; minimum observed value=0, maximum observed

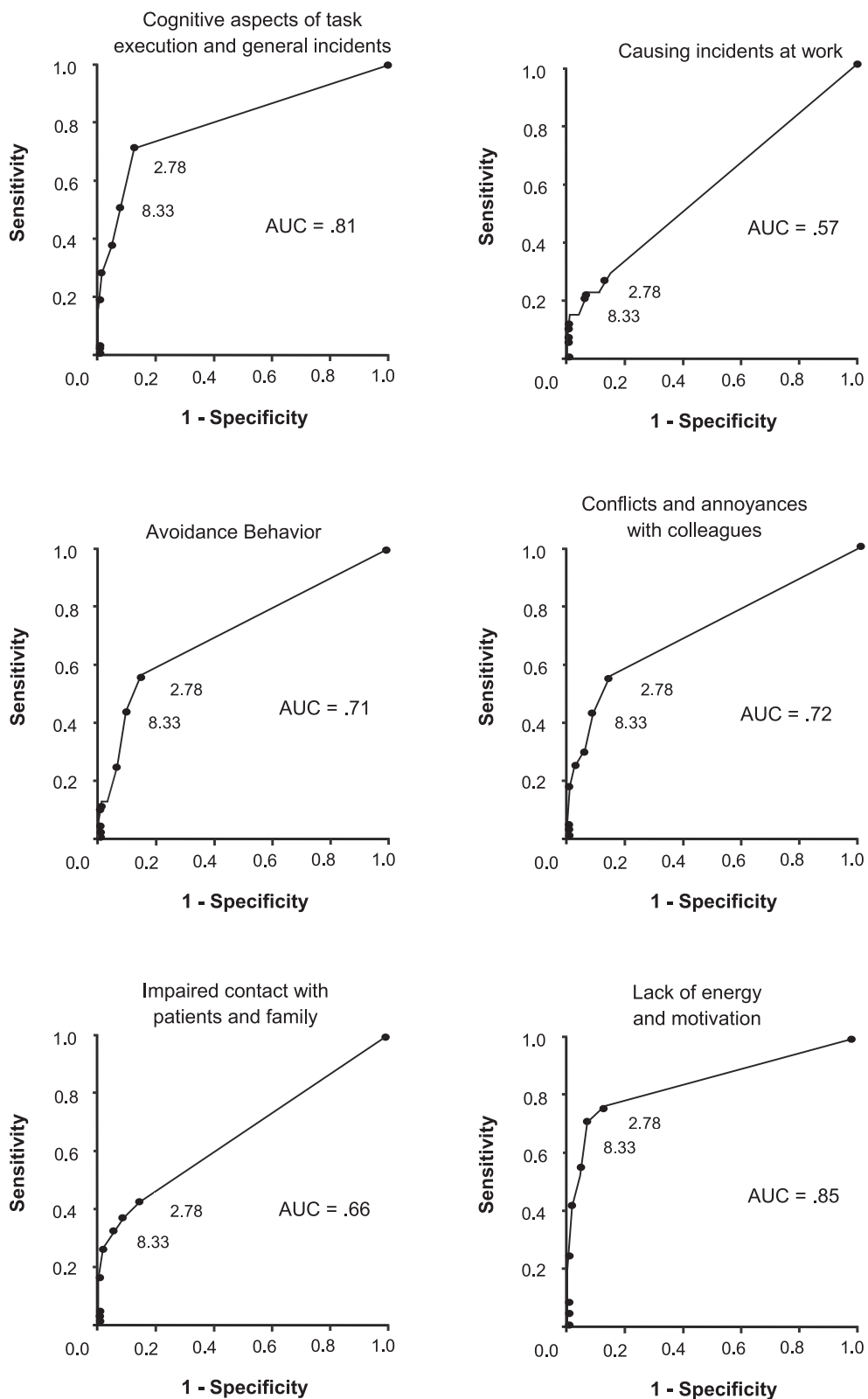


Fig. 1. ROC-Curves

value=81); conflicts and irritations with colleagues (mean =8.5, SD=12.4; minimum observed value=0, maximum observed value=61); impaired contact with patients and

their family (mean=6.9, SD=8.8; minimum observed value=0, maximum observed value=42); and lack of energy and motivation (Mean=10.5, SD=13.7; minimum

Table 1. Predictive properties 6-item work functioning screener (WFS-H) for various aspects of work functioning

Work functioning domain (NWFQ)	Sensitivity	Specificity	AUC	NPV	PPV
1.) Cognitive aspects of task execution and general incidents	71%	87.44%	.79***	97%	36%
2.) Avoidance behavior	56%	85%	.71**	96%	22%
3.) Conflicts and irritations with colleagues	56%	85%	.71**	96%	21%
4.) Impaired contact with patients and family	42%	86%	.64*	94%	22%
5.) Lack of energy and motivation	76%	87%	.82***	98%	32%

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. AUC=Area-under-the-curve, NPV=negative predictive value, PPV=Positive predictive value. The AUCs presented resulted from ROC-analyses conducted to examine whether the WFS-H with dichotomous values (i.e., negative and positive) could serve as a test for the work functioning domains.

Table 2. Correlations WFS-H scores with scores comparator measures

Comparator measure (n=236)	Screener Work Functioning (WFS-H) <i>r</i>
1.) Productivity (EWPS)	.51***
2.) General Health (GHQ)	.48***
3.) Distress (4DKL)	.47***

Note. *** $p < .001$. n=236 due to missing values.

observed value=0, maximum observed value=73). In terms of these work functioning domains, the ROC analyses further suggested a cut-off point between 2.8 and 8.3 on a 0-100 scale for the WFS-H (Fig. 1). The middle value of five and a half was chosen as the definite cut-off point. This cut-off point corresponds with a raw WFS-H score of 1 ($(1 \cdot 100)/18=5.5$), while the cut-off point 2.7 did not translate to a raw WFS-H score that respondents can actually score. The cut-off point 8.3 had less optimal balance between sensitivity and specificity than the chosen cut-off point (5.5). Further, a score below 5.5 corresponds with selection of the answer option "not at all" for all WFS-H items by individual respondents, and thus represents a clear indication of ability to meet job demands that are difficult to complete for individuals with mental health problems. In addition, the WFS-H contains items of varying conceptual content since work functioning is a multi-domain concept. Respondents who answer positive to only one or a few of the items of the WFS-H thus may experience impaired work functioning due to mental health problems in specific work functioning domains, and this would be overlooked in case the cut-off point would exceed the value of five and a half (5.5). Thus, for multiple reasons 5.5 was chosen as the definite cut-off point, and a WFS-H score below 5.5 indicates a clear negative test result, and a WFS-H score between 5.5 and 100 indicates a positive WFS-H test result. Subsequently,

the predictive value of the WFS-H (Mean=2.7, SD=7.4; minimum observed value=0, maximum observed value=61) was examined. For an overview, the indicators of the predictive value of the WFS-H are displayed in Table 1. A negative WFS-H test score suggests that decreased work functioning due to mental health problems is highly unlikely in terms of cognitive task execution and general incidents (negative predictive value: 97%; positive predictive value: 36%;) and energy and motivation for the job (negative predictive value: 98%; positive predictive value: 32%). A negative WFS-H screener score also is highly indicative of absence of avoidance behavior at work (negative predictive value: 96%; positive predictive value: 22%) and absence of conflicts and annoyances with colleagues (negative predictive value: 96%; positive predictive value: 21%). Finally, a negative WFS-H screener score moderately suggests that the contact with patients and their family members (negative predictive value: 94%; positive predictive value: 22%) is sufficient and not hindered by mental health problems. Thus, the newly developed brief work functioning screener (WFS-H) has value for detecting healthcare workers without work functioning problems. Table 2 shows the associations between the brief work functioning screener (WFS-H) on the one hand, and the measures used for testing the convergent construct validity of the brief work functioning screener on the other hand. As predicted, the WFS-H correlates ($p < .001$) moderately with productivity ($r=.51$), general health ($r=.48$), and distress ($r=.47$). The work functioning screener therefore has good convergent construct validity.

Discussion

The aim of the current research was to examine the predictive value and convergent construct validity of the WFS-H. It was reasoned that a negative WFS-H test score would indicate sufficient ability to (continue to) fulfill work demands that are difficult to complete for individu-

Appendix. Work Functioning Screener-Healthcare (WFS-H)

Answer options: 0='not at all', 1='slightly agree', 2='agree', 3='totally agree'

- 1) My work functioning is worse than before
- 2) Without the help of my colleagues I would not function adequately
- 3) I make mistakes more often
- 4) I take risks more often in my job than before
- 5) I retire to a less hectic place more than before
- 6) For a while now I have had less interest for my work activities

Note. The screener sumscore ranges from 0 to 18. A 0 to 100 score can be calculated with the following formula: $\text{Score}_{0 \text{ to } 100} = ((\text{raw sumscore WFS-H} * 100) / 18)$.

als with mental health problems. Healthcare workers who had a negative WFS-H test score predominantly were found to be workers who reported work functioning unhindered by mental health problems. Further, as predicted, the WFS-H test score correlated moderately with productivity, generic mental health, and distress levels of workers. In view of the results, it was concluded that the WFS-H has good predictive and good convergent construct validity.

Implications of the research

It is important to evaluate and sustain the work functioning of healthcare workers because impaired work functioning of healthcare workers is detrimental to the recovery process and quality of life of recipients of medical care services. The current research represents a contribution to the literature, because it developed and validated a brief screening instrument (WFS-H) for gaining a preliminary insight into the work functioning of healthcare workers. Occupational health professionals can use this instrument to determine whether healthcare workers sufficiently meet job demands that are difficult to complete for individuals with mental health problems. The instrument also helps in identifying healthcare workers who have a potential need for care and/or work support. Healthcare workers with a positive test score should be asked for additional information on their work functioning. An instrument such as the NWFQ can be used for this. The current research thus suggests a two-step procedure for evaluating the work functioning of healthcare workers. It is most efficient to first evaluate whether healthcare workers sufficiently meet their work demands using the WFS-H, and then administer longer work functioning questionnaires (e.g., NWFQ) only among healthcare workers with a positive WFS-H test score.

Limitations and suggestions for further research

The current research only focused on the predictive value and convergent construct validity of the WFS-H. Furthermore, only one sample was used in the research, consisting of only nurses as healthcare workers. It is therefore important to acknowledge that the findings of

the current research are not conclusive. Future studies should examine whether the current findings can be replicated, and researchers should further examine the validity of the WFS-H (e.g., divergent validity, cross-cultural construct validity, etc.). Users should be aware of this when using the WFS-H for examining the work functioning of healthcare workers. Another limitation of the research lies in the fact that the "gold standard" instrument in this research (NWFQ) is a self-report instrument. Objective indicators of work functioning, such as supervisor plus co-worker ratings of work performance and number of days absent from work, may better reflect the actual work functioning of healthcare workers, and thus be better "gold standard" measures. At the same time, the NWFQ has good psychometric properties¹⁵⁾ and is an instrument that researchers frequently use for recording the work functioning of healthcare workers^{16,17)}. Thus, for now, it is a recommendation to examine whether the WFS-H also has predictive value for work functioning scores obtained with objective indicators. Further, researchers can examine whether the recommended two-step procedure for evaluating the work functioning of healthcare workers is to the satisfaction of occupational health professionals. Also, researchers can use the WFS-H in studies on health surveillance, and, for instance, examine whether the WFS-H test score correlates with healthy lifestyle choices of healthcare workers.

In conclusion

The currently developed work functioning screener can assist occupational health professionals in guiding healthcare workers and ensuring continued quality of care for those receiving medical aid from healthcare workers.

References

- 1) Noben C, Smit F, Nieuwenhuijsen K, et al. Comparative cost-effectiveness of two interventions to promote work functioning by targeting mental health complaints among nurses: Pragmatic cluster randomised trial. *Int J Nurs Stud* 2014; 51: 1321-1331.
- 2) Weinberg A, Creed F. Stress and psychiatric disorder in healthcare professionals and hospital staff. *Lancet* 2000; 355:

- 533-537.
- 3) Gärtner FR, Nieuwenhuijsen K, Van Dijk FJ, Sluiter JK. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010; 47: 1047-1061.
 - 4) Nieuwenhuijsen K, Franche RL, Van Dijk FJ. Work functioning measurement: Tools for occupational mental health research. *J Occup Environ Med* 2010; 52: 778-790.
 - 5) Blozik E, Scherer M, Lacruz ME, Ladwig KH; KORA Study Group. Diagnostic utility of a one-item question to screen for depressive disorders: Results from the KORA F3 study. *BMC Fam Pract* 2013; 14: 198-205.
 - 6) Cairney J, Veldhuizen S, Wade TJ, Kurdyak P, Streiner DL. Evaluation of 2 measures of psychological distress as screeners for depression in the general population. *Can J Psychiatry* 2007; 52: 111-120.
 - 7) Jacobsen PB, Donovan KA, Trask PC, Fleishman SB, Zabora J, Baker F. Screening for psychologic distress in ambulatory cancer patients: A multicenter evaluation of the distress thermometer. *Cancer* 2005; 103: 1494-1502.
 - 8) Gore KL, Enger CC, Freed MC, Liu X, Armstrong DW 3rd. Test of a single-item posttraumatic stress disorder screener in a military primary care setting. *Gen Hosp Psychiatry* 2008; 30: 391-397.
 - 9) Gärtner FR, Nieuwenhuijsen K, Van Dijk FJ, Sluiter JK. Impaired work functioning due to common mental disorders in nurses and allied health professionals: The nurses work functioning questionnaire. *Int Arch Occup Environ Health* 2012; 85: 125-138.
 - 10) Stanton JM, Sinar EF, Balzer WK, Smith PC. Issues and strategies for reducing the length of self-report scales. *Pers Psychol* 2002; 55: 167-194.
 - 11) Henkel V, Mergl R, Coyne JC, Kohnen R, Möller HJ, Hegerl U. Screening for depression in primary care: Will one or two items suffice? *Eur Arch Psychiatry Clin Neurosci* 2004; 254: 215-223.
 - 12) Lerner D, Amick BC III, Rogers WH, Malspeis S, Bungay KM, Cynn D. The work limitations questionnaire. *Med Care* 2001; 39: 72-85.
 - 13) Amick BC III, Habeck RV, Ossmann J, Fossel AH, Keller R, Katz JN. Predictors of successful work role functioning after carpal tunnel release surgery. *J Occup Environ Med* 2004; 46: 490-500.
 - 14) Clarke SP, Douglas MS, Aiken LH. Effects of hospital staffing and organizational climate on needlestick injuries to nurses. *Am J Public Health* 2002; 92: 1115-1119.
 - 15) Gärtner FR, Nieuwenhuijsen K, Van Dijk FJ, Sluiter JK. Psychometric properties of the Nurses Work Functioning Questionnaire (NWFQ). *PLoS One* 2011; 11: e26565.
 - 16) Gärtner FR, Ketelaar SM, Smeets O, et al. The Mental Vitality[®] Work study: Design of a randomized controlled trial on the effect of a workers' health surveillance mental module for nurses and allied health professionals. *BMC Public Health* 2011; 11: 290.
 - 17) Ketelaar SM, Nieuwenhuijsen K, Gärtner FR, Bolier L, Smeets O, Sluiter JK. Effect of an e-mental health approach to workers' health surveillance versus control group on work functioning of hospital employees: A cluster-RCT. *PLoS One* 2013; 9: 1-11.
 - 18) Terwee CB, Bot SDM, De Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42.
 - 19) Letvak S, Ruhm CJ, McCoy T. Depression in hospital-employed nurses. *Clin Nurse Spec* 2012; 26: 177-182.
 - 20) Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull* 1997; 33: 13-16.
 - 21) Goldberg D, Bridges K, Duncan-Jones P, Grayson D. Detecting anxiety and depression in general medical settings. *BMJ* 1988; 297: 897-899.
 - 22) Jackson C. The general health questionnaire. *Occup Med* 2007; 57: 79.
 - 23) Terluin B. De Vierdimensionele Klachtenlijst (4DKL) in de huisartspraktijk [The Four Dimensional Symptom Questionnaire (4DSQ)]. *De Psycholoog* 1988; 33: 18-24.
 - 24) Terluin B, van Marwijk HW, Ader HJ, et al. 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: 34.
 - 25) Streiner DL, Cairney JC. What's under the ROC? An introduction to receiver operating characteristics curves. *Can J Psychiatry* 2007; 52: 121-128.
 - 26) Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007; 26: 2170-2183.