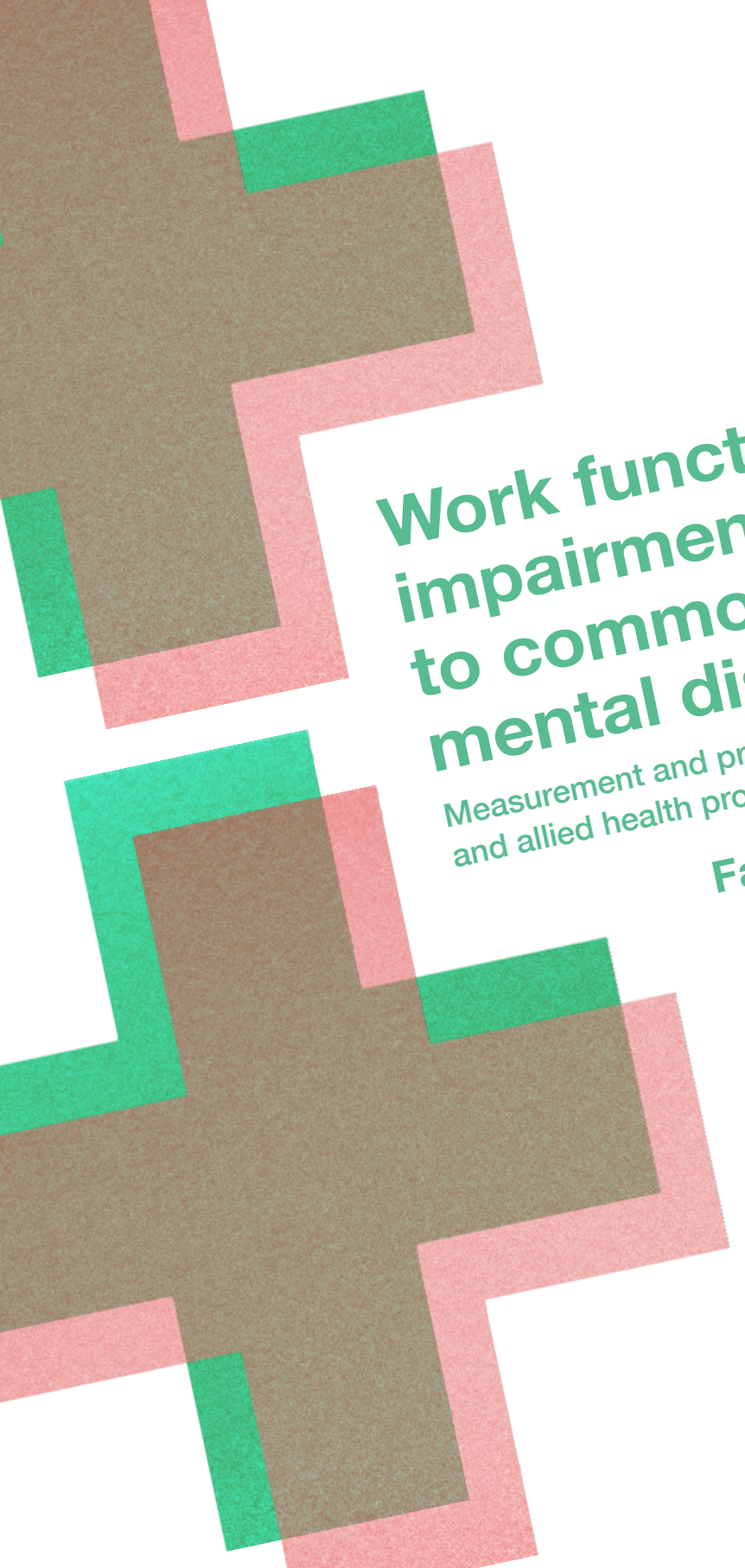
Work functioning impairments due to common mental disorders: measurement and prevention in nurses and allied health professionals

Gärtner, F.R.



Work functioning impairments due to common mental disorders

Measurement and prevention in nurses and allied health professionals

Fania R. Gärtner

Work functioning impairments due to common mental disorders

Measurement and prevention in nurses and allied health professionals

Fania R. Gärtner

All studies described in this thesis were carried out at the Academic Medical Center, University of Amsterdam, Department: Coronel Institute of Occupational Health, Amsterdam, the Netherlands.

The studies in this thesis were financially supported by the Dutch Foundation Institute Gak and partly by the Netherlands Organization for Health Research and Development (ZonMw).

Cover design: Dennis Keeskamp: www.spulvankeeskamp.com
Printing: CPI – Wöhrmann Print Service
ISBN: 978-94-91043-06-2

Please cite as: Gärtner F.R. (2012). *Work functioning impairments due to common mental disorders. Measurement and prevention in nurses and allied health professionals* (Doctoral Dissertation). Amsterdam: University of Amsterdam.

Work functioning impairments due to common mental disorders

Measurement and prevention in nurses and allied health professionals

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 7 september 2012, te 10:00 uur

door

Fania Ruth Gärtner
geboren te Berlijn, Duitsland

Promotiecommissie

Promotor:	Prof. dr. F.J.H. van Dijk
Copromotores:	Dr. K. Nieuwenhuijsen
	Dr. J.K. Sluiter
Overige leden:	Prof. dr. J.R. Anema
	Dr. U. Bültmann
	Prof. dr. P.M.G Emmelkamp
	Prof. dr. R.J. de Haan
	Prof. dr. H.C.W. de Vet
	Prof. dr. H.P.C.M. van Weert

Faculteit der Geneeskunde

CONTENTS

1

General introduction

Common mental disorders and work

Mental disorders, such as adjustment, mood, and anxiety disorders, are highly prevalent in most industrialized countries. For the general working population in Europe, the 12-month prevalence for having any depressive, anxiety, or alcohol-related disorder is 9%.[1] The prevalence of milder symptoms is even higher. Based on two Dutch studies, the prevalence for distress, mild depression, or mild anxiety ranges from 16% to 22% in adult men and from 19% to 26% in adult women.[2 3] This group of mild and moderate severe mental disorders is conceptualized as common mental disorders.[4 5] For the studies in this thesis, in accordance with Glozier,[4] Watanabe and colleagues,[6] and based on their prevalence[5 7 8] adjustment, depressive, and anxiety disorders, as well as substance abuse and dependence are included in the umbrella term of common mental disorders.

Common mental disorders (CMDs) are known to have a high burden of disease for patients, impairing their daily functioning and quality of life.[9] Aside from the burden on one's personal life, CMDs also impact one's participation and functioning at work.[10-14] The World Health Organization (WHO) expects depression to be the leading cause of absenteeism in industrialized countries in 2020.[15] Absenteeism in relation to CMDs, its causes, costs, and related interventions has been studied extensively in the past. However, CMDs can also have substantial impact on the work functioning of workers attending work while suffering from mental health complaints, which is often referred to as presenteeism.[12 13 16-23] Stewart and colleagues[24] found that 81% of the costs for lost productive time of workers with depression were due to impaired performance while at work. In some occupations, such as healthcare services, CMD-related work functioning impairments can have serious consequences, as they can cause incidents with risks for the health of workers and patients.[25]

Common mental disorders and the work of nurses and allied health professionals

In the nursing profession impaired work functioning due to CMDs is of special importance for three reasons. First, in the work of nurses – including surgical nurses and anesthetic nurses – the disposition to be absent from work is lower, which leads to more presenteeism compared to other professions.[16] Second, impairments in the work functioning of nurses with CMDs can have serious implications for the health and safety of themselves and their patients. Previous studies in nurses have shown that workers with poor mental health experience significantly more medical accidents.[25 26] Third, nurses have a high incidence of CMDs. The relative risk of depression is highest for nurses, RR = 3.5, 95% CI (1.3, 9.6), than for other human service workers and other healthcare workers.[27] A possible explanation for this high prevalence may be found in work characteristics, which include high job demands, high emotional demands, a lack of autonomy, and insufficient social support.[28 29] These are known risk factors for developing mental health complaints.[30]

Next to nurses, allied health professionals, such as physiotherapists and radiotherapists, form another large group of workers in the hospital setting. To our knowledge, no research

findings on CMDs and associated work functioning impairments have been published about this group to date. Recognizing differences among the various occupations, departments and local conditions, nurses and allied health professionals have many work demands and conditions in common. Therefore, it is expected that the magnitude of CMDs and related work functioning problems is similar for these two groups of workers. In this thesis, I study both nurses and allied health professionals, and from here on I refer to them as one occupational group.

Until now, detailed knowledge on the exact impact of CMDs on work functioning is scarce, as most previous research on presenteeism expressed impaired work functioning quantitatively only, e.g., in terms of lost days,[24 31] days attending work when feeling ill,[17] or work days that require extra effort to function well when suffering from ill health.[32] Those studies provide insight into the magnitude of impaired work functioning but do not reflect the nature of impaired work functioning. Few researchers have identified specific aspects of work that are impaired due to mental health complaints. Almost three decades ago, Motowidlo[33] differentiated seven aspects of work performance in nurses that are negatively influenced by depressed feelings at work including the quality of patient care and cognitive, intra- and interpersonal aspects, as rated by supervisors and colleagues. In the first decade of the 21st century, more studies were published that distinguished specific aspects of work functioning for the working population in general. Lerner and colleagues[34] showed that depression affects mental performance, interpersonal tasks, time management, and overall performance. Wang and colleagues[35] differentiated two aspects of work performance that are impaired due to depression as follows: task focus or concentration and productivity (quality, speed, and efficiency of task completion).

If we want to reduce the burden and risks associated with work functioning impairments due to CMDs, we might profit from further insight into which different aspects of work functioning are affected by CMDs. Instruments for detection and monitoring of impaired work functioning can then be developed, which would enable future research on the onset of and recovery from the impairments. Additionally, intervention strategies might be developed to intervene purposefully in the identified aspects of impaired work functioning.

Conceptual model of work functioning

The elements of work functioning

No scientific consensus exists regarding what exactly constitutes work functioning. Various lines of research approach this topic differently. Based on the literature from the occupational health field and of work and organizational psychology, the following conceptual model of work functioning was developed (see Figure 1). First, two key elements of work functioning are distinguished in correspondence with the performance literature: task versus contextual aspects of work functioning.[36 37] The task aspect includes all activities performed by a worker

to accomplish the core work tasks. The contextual aspect relates to all behavior that supports the organization and the social and psychological environment in which the tasks are executed,[37] including interpersonal interactions and helping behavior.

Both task and contextual work functioning can be further specified, which is performed in the studies of this thesis with a focus on nurses and allied health professionals. To guide the specification of various aspects of work functioning, we formulated four dimensions that can be applied to both task and contextual performance. The process of working is the first dimension distinguished, which can be described as the behavior exerted to achieve work outcomes.[38 39] In other words, it regards evaluation of what a person is doing at work and how the person is completing his or her tasks.[39] The second and third dimensions that we included follow the distinction between the quantity and the quality of work outcomes. In our opinion, it is not sufficient to focus only on quantity outcomes, e.g., days or hours of lost productivity, as is performed in many studies.[24 40] For a comprehensive examination of work functioning, the quality of work outcomes is also relevant. In previous research, the quality of work is not expressed in economical terms but rather in terms of increased risk for incidents or near-misses associated with underlying health impairments.[25 41 42] The fourth dimension of work functioning in this conceptual model is the extra effort required by a worker to function well at work despite mental health complaints. It was found that workers with depression achieve normal productivity on some days but sometimes at the cost of extreme personal effort.[32]

The context of work functioning

Actual work functioning always takes place in a concrete context, which has to be regarded when studying impairments in practice. Aside from the health condition of a worker, the forces working within this context are environmental and personal factors (see Figure 1). This idea is in line with the International Classification of Functioning Disability and Health (ICF) of the WHO.[43] In the ICF scheme, functioning and participation are characterized as the result of a complex relationship between the individual's health condition, (other) personal factors, and environmental factors (physical and social) that represent the circumstances in which the individual lives. These three components can positively and negatively influence participation. Due to these components, the same health conditions impact the functioning and disability of individuals differently, because different environmental and personal factors influence this relationship.

When applying the ICF scheme to functioning at work, work tasks and work conditions can be regarded as the environmental factors. It can be expected that some work demands and conditions lead to different or more limitations in work functioning than others. Therefore, differences in work functioning impairments can be expected for different occupational groups. This idea is elaborated in a systematic review about factors associated with work functioning in depressed workers by Lagerveld and colleagues.[11] One study found that in jobs involving proficiency in exercising judgment and communication tasks – such as the jobs of

healthcare workers – more limitations in work functioning were found in depressed workers than in jobs that do not require these skills.[34] Therefore, a job-specific approach to work functioning impairments due to CMDs might be useful for research and in practice.

Personal factors influencing work functioning and its impairment include a variety of aspects of the individual's background. For instance, personality characteristics are known to be associated with a high level of work functioning impairment, such as higher neuroticism, more external locus of control, and lower self-esteem.[44] Other characteristics may also be relevant, such as the professional expertise and personal priorities of a worker, e.g., which tasks are regarded as most favorable or important.

In the conceptualization of work functioning for this thesis, work functioning is presented as one concept including different dimensions. It is not assumed that all of these dimensions will be equally affected by CMDs. Rather, the aspects of work that are affected may differ from one person to another. For example, in one worker with depression, the quality of work will suffer, while in another worker with depression the quantity of work will decrease, whereas the quality of work might be unaffected. This assumption supports the need for a multidimensional approach of work functioning impairments due to CMDs.

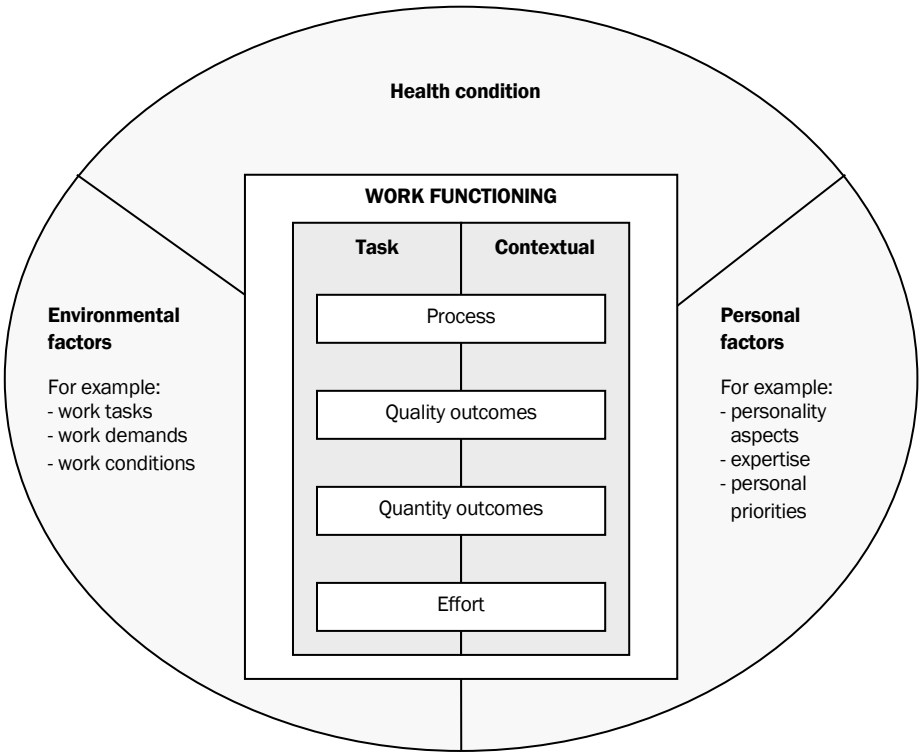


Figure 1. Conceptual model of work functioning.

The measurement of work functioning impairments due to common mental disorders

Based on the description above, three requirements for instruments to assess work functioning and its impairments can be defined as follows: disease specificity, job specificity, and multidimensionality. A fourth aspect we regard as necessary for the measurement instrument is its suitability for the entire occupational group. It is assumed that even mild CMDs can result in impaired work functioning, even though the worker might not be aware of the presence of mental health problems and their consequences.[45 46] Thus, if aiming to use a work functioning questionnaire for evaluative purposes or possibly also for the identification of individuals in need of interventions to improve work functioning, measurement instruments should be able to be administered in subjects with unknown or mild mental health complaints. Therefore, the items included in these instruments should not refer to any specific mental health or emotional problems.

Recently, two systematic reviews were published on existing work functioning questionnaires, their application in a CMD working population, and the associated clinimetric qualities.[47 48] (Note: when referring to measurement properties the term *clinimetric* quality is used in this thesis, except for chapter 4, where the term *psychometric* properties is applied.) Twenty-one work functioning measures have been found in the literature;[48] however, the requirements for measurements of work functioning impairment formulated above are not met in most of the existing questionnaires. First, regarding the disease specificity of existing instruments, a review by Abma and colleagues[47] identified five instruments, of which the clinimetric quality was studied in a CMD population. Only one questionnaire, the Lam Employment Absence and Productivity Scale (LEAPS), was developed specifically for this patient group.[47] Nieuwenhuijsen and colleagues identified 11 instruments that have been applied in this patient group; however, data on the clinimetric properties established in CMD patients were missing for most of these instruments.[48] Second, no job-specific instrument has been found. Third, most of the existing work functioning questionnaires do not meet the demand for multidimensionality. In the review by Abma and colleagues, only two of the five questionnaires assessed multiple dimensions: the Work Limitation Questionnaire (WLQ) comprises four subscales, and the Lam Employment Absence and Productivity Scale (LEAPS)[49] has two subscales. Fourth, many of the existing work functioning scales, e.g., the Work Limitations Questionnaire (WLQ)[50] and the Stanford Presenteeism Scale (SPS),[51] refer explicitly to health problems in their items before the description of a specific functioning problem is given. Therefore, these questionnaires are less suitable for administration in an apparently healthy working population and thus, also for detecting new cases of workers with impaired work functioning due to mild and often undetected mental disorders.[48]

Overall, it is concluded that there are sound arguments for the development of a new work functioning instrument that fulfills the requirements described above.

Seeking help for mental health complaints

Although effective care for mental health complaints is available, mental health care is often not utilized by workers with CMDs.[52-54] Regarding occupational health care in the Netherlands, it is known that workers in most cases make use of its service only after they are sick-listed. Late or no help-seeking for mental health complaints is a well-known phenomenon in primary care.[52 53] A qualitative study in human service workers with burnout revealed that help-seeking often happens only after a breaking point.[54] These breaking points can be incidents or conflicts at work. Two reasons for absent or late help seeking are known. Firstly, a lack of recognition of one's own mental health problem seems to be an obstacle to help seeking. In a study by Lexis and colleagues,[55] 43% of workers who had mild to severe mental health complaints, as indicated by an online screening, reported not to perceive mental health complaints. Moreover, in a study across six European countries, only 33% of adults with a mental health disorder perceived the need for mental health care.[56] Secondly, it has been found that attitudinal aspects, such as a fear of stigmatization or lack of trust in health care are barriers to active help seeking.[57 58] The prevailing negative attitude regarding mental health care is illustrated in a study by Ten Have and colleagues,[59] in which 46% of randomly selected adults in the Netherlands regard the effectiveness of professional help for mental health complaints as "worse" or "equal to" no treatment. Furthermore, in that study, 35% of the Dutch adults stated that they would "probably not go" or "definitely not go" to a professional in the case of serious emotional problems. Healthcare workers in particular seem to experience barriers when seeking treatment for their own (mental) health complaints.[60 61] This group finds it hard to seek help, as they are used to providing care instead of receiving care.[54]

When mental health complaints go untreated for a long period of time, they can become worse. Thus, the later patients receive help, the more difficult successful treatment becomes, which prolongs the duration of illness.[62] Preventive actions might be useful to detect workers with early signs of CMDs and to encourage active and early help-seeking behavior in this group. This form of preventive action is classified as indicated prevention.[63] In relation to the content of this thesis, indicated prevention implies the screening of preliminary or mild symptoms of CMDs and early signs for impaired work functioning to prevent the onset or progression of CMDs or work functioning impairments. If such prevention can be offered successfully within a work setting, timely help for workers with CMDs can be provided. Work functioning that is impaired to the extent that workers cause serious incidents or that they must call in sick, can thus be prevented. Within the occupational healthcare setting, one well-developed strategy for preventive actions is workers' health surveillance (WHS).[64 65] WHS is a prevention strategy that aims at the early detection of negative health effects of work and of the inability to work to enable timely interventions.[65] Regarding mental health aspects, a WHS mental module may be a successful preventive strategy to prevent the deterioration of CMDs and to prevent impairments in work functioning and work disability in the healthcare sector.

Workers' health surveillance

WHS is a periodical assessment of employees' health and work, e.g., carried out every two to four years. Usually, WHS is voluntary for employees, except for workers in high-demand jobs, such as firefighters.[66] In 1998, the International Labour Organization (ILO) formulated criteria for the use of WHS.[64] The four criteria that have to be met are: need, relevance, scientific validity, and effectiveness.

Despite the international recommendations of the ILO, the design of WHS differs among countries. In the Netherlands, a policy guideline on how to conduct WHS was published in 2005 by the Netherlands Society of Occupational Medicine.[66] In the Netherlands, unlike some other countries, WHS does not include the assessment of risk factors, as these factors are included in a separate prevention strategy for risk assessment called the Risk Inventory and Evaluation (RI&E). The core aims of WHS according to the Dutch guidelines are the monitoring and improvement of any health problem in relation to the work and the monitoring and improvement of work functioning problems of individual workers along with the detection of occupational diseases. The focus on work functioning as it relates to health is explicit in the guidelines. According to the aim of monitoring and improvement, WHS comprises two key elements: first a screening, to detect employees with (incipient) health or work functioning problems; second, the provision of interventions to improve health and work functioning and to reduce the risk of deteriorations in health status.

Historically and also in the most recently developed WHSs, in the Netherlands WHS aims to target the health hazards and demands present in specific occupations.[67 68] More recently, this approach has been referred to as *job-specific* WHS.[67] With this job-specific approach, WHS are regarded as more likely to be needed, relevant, valid, and effective.[67] Therefore, for a newly developed WHS mental module, the screenings and interventions are expected to closely reflect the demands and risks of the work tasks of specific occupations.

Although the use and application of WHS is increasing for various occupations and health effects, little is known about WHS targeting mental health effects. In a recent literature review by Plat et al.[69] on WHS among military and emergency service personnel, only three of the 24 studies included mental health aspects: one in police personnel,[70] one in rescue and recovery workers,[71] and one in soldiers.[72] It has been argued that a focus on mental health should be included in occupations with a high risk of developing CMDs.[15] As aforementioned, these risks are also present in the work of nurses and allied health professionals.[28 29] A WHS mental module for nurses and allied health professionals might stimulate insight into one's own mental health state and work functioning, through which help-seeking behavior might be encouraged. Additionally, active help seeking among workers with CMDs or work functioning problems might be facilitated by the invitation to consult with an occupational physician. Help seeking might in turn lead to an improvement in work functioning impairments and mental health. However, as far as it is known, a WHS for mental health and its consequences for the

work of nurses or other healthcare service professionals has not yet been scientifically developed and evaluated.

A mental module for workers' health surveillance among nurses and allied health professionals

A new WHS mental module for nurses and allied health professionals in the Netherlands should follow the combined requirements formulated by the ILO and the Dutch guidelines. First, both the screening and the intervention part should focus on health aspects as well as on the work functioning of employees. Second, the WHS mental module should reflect the hazards and demands of the work of nurses and allied health professionals and thus be job-specific. Third, the WHS mental module should be scientifically evaluated for its effectiveness.

Below, the development of the WHS mental module that is to be evaluated is further described. The description is presented separately for the two elements, screening and intervention.

Workers' health surveillance mental module: screening

For the screening part of the WHS mental module, various forms are conceivable. Self-report screening questionnaires are regarded as useful to enhance feasibility and to guarantee the confidential nature of screening. Self-report questionnaires are easy to administer in a reasonable amount of time compared to diagnostic interviews. Moreover, these questionnaires do not depend on third persons and, therefore, no supervisors or co-workers need to be involved in the assessment. As the WHS mental module will be offered at the level of the department or organization, large-scale screening must be feasible. Offering screening questionnaires online enables a large group of workers to utilize the WHS mental module.[73]

The screening part ought to include screening for mental health complaints as well as for impaired work functioning. The screening for mental health aims to detect workers with mild to severe mental health complaints related to the work of healthcare workers. Therefore, there are separate questionnaires used for adjustment, fatigue, depression, and anxiety disorders, as well as for risky drinking behavior. To this end, there are several validated questionnaires that are also suited for application in the working population.[74-77]

The screening for work functioning impairments should reflect all relevant facets of work exertion possibly impaired by CMDs. A job-specific questionnaire is regarded as advantageous for three reasons. First, items that give concrete examples of tasks ask for fewer interpretations and therefore make self-reporting easier. Second, the concrete examples of everyday work might provide insight into one's own functioning and therefore stimulate the recognition of possible impairments. Third, screening results that present concrete examples of work tasks that may be impaired provide valuable input for interventions by the occupational physician and also for adaptations in the work or behavior initiated by the worker

him- or herself. The job-specific questionnaire developed as part of this thesis will therefore be used.

Workers' health surveillance mental module: intervention

This screening strategy and the feedback of its results are supposed to stimulate reflection on one's own mental health. They are therefore expected to enhance help-seeking behavior in workers who are screened as positive for work functioning impairments and/or mental health complaints. To further encourage help-seeking behavior, an invitation for a preventive consultation with the occupational physician is included in this WHS mental module for workers who are screened as positive. It is presumed that visiting a care-giver (perchance the occupational physician) enhances the work functioning and mental health, by the interventions that would be initiated.

The care as usual provided by occupational physicians in the Netherlands for workers with CMDs follows the evidence-based guidelines of the Netherlands Society of Occupational Medicine.[78] [79] This care can be considered effective; however, it mainly focuses on the guidance of workers on sick-leave.[78] [80] Little is known on the effectiveness of care provided by occupational physicians in a preventive setting. Therefore, a protocol will be developed for the design of the WHS preventive consultation. This protocol will be based on the Dutch guidelines for the care provided by occupational physicians and adapted to the specific context of an open consultation. These consultations are taken on the employee's initiative, which is in contrast to mandatory consultations for sick-listed employees. One difference from regular open consultations is that employees in the WHS setting do not always perceive the need for help but instead just follow upon the invitation to attend a consultation. The extent to which worker recognize their screening results therefore requires special attention in the preventive consultation.

THESIS AIMS AND RESEARCH QUESTIONS

Two main aims are formulated for this thesis.

Aim I

The first aim of this thesis is to develop a job-specific instrument to assess work functioning impairments due to common mental disorders in nurses and allied health professionals. This instrument is aimed to be multidimensional and in the form of a self-report questionnaire. The new instrument will be evaluated for its reliability, its validity, and the interpretability of change scores.

Research question i: What aspects of work functioning can be impaired due to CMDs in nurses and allied health professionals? (Chapters 2 & 3)

Research question ii: What are the content validity, factorial structure, and internal consistency of a newly developed questionnaire for work functioning impairments due to common mental disorders in nurses and allied health professionals – the Nurses Work Functioning Questionnaire (NWFQ)? (Chapter 3)

Research question iii: How is the clinimetric quality of the NWFQ evaluated in terms of the reproducibility, construct validity, and interpretability of change? (Chapters 4 & 5)

Aim II

The second aim concerns the evaluation of a newly developed workers' health surveillance (WHS) mental module for nurses and allied health professionals. The WHS mental module aims to identify workers with mild to moderate severe symptoms of CMDs and/or early signs of impaired work functioning in an online screening. Workers, who are screened as positive for either mental health complaints or work functioning impairments, or both, are offered a consultation with the occupational physician to receive advice on appropriate interventions and care providers. The hypothesis is that workers who take part in the WHS mental module will show more help-seeking behavior compared to a control group. Also, in that group of workers an improvement of work functioning impairments and mental health complaints is expected.

Research question iv: Is a workers' health surveillance mental module an effective strategy to stimulate help-seeking behavior and to improve work functioning and mental health in nurses and allied health professionals with mild to moderate severe mental health complaints and/or work functioning impairments due to common mental disorders, compared to a control group? (Chapters 6 & 7)

Research question v: How is the workers' health surveillance mental module evaluated in terms of response, compliance, adherence, and perspectives on the workers' health surveillance mental module among the workers and the occupational physicians? (Chapter 8)

THESIS OUTLINE

This thesis comprises two main parts: first, the development and evaluation of a work functioning questionnaire related to common mental disorders; second, the evaluation of a workers' health surveillance mental module. In **Chapter 2**, a systematic literature review is presented which aimed to identify aspects of work functioning among nurses and allied health professionals that are impaired due to common mental disorders. **Chapter 3** focuses on the development of a questionnaire to assess impaired work functioning in nurses and allied health professionals with common mental disorders, called the Nurses Work Functioning Questionnaire (NWFQ). In addition to the literature review, data from a focus group study were used to develop the questionnaire. Furthermore, the content validity and factorial structure are analyzed and discussed in this chapter. In Chapters 4 and 5, the clinimetric qualities of the NWFQ are evaluated. **Chapter 4** focuses on the reproducibility and the construct validity of the NWFQ. The reproducibility analysis includes information on the agreement of the questionnaire (the extent of differences between repeated measures) and the reliability (the ability of the questionnaire to distinguish between persons despite measurement error). The construct validity provides information on whether the questionnaire measures what it intends to measure. **Chapter 5** concerns the interpretability of change of the NWFQ. In the study described in this chapter, the smallest detectable change and the minimal important change values for the NWFQ total scale and its subscales were identified. In Chapters 6, 7 and 8, evaluation of the WHS mental module for nurses and allied health professionals is addressed. **Chapter 6** describes the design of a cluster-randomized trial in which the effectiveness of an intervention strategy for a WHS mental module is studied and compared to a control group that did not receive the WHS mental module, including screening results and consultation. **Chapter 7** addresses the effect of the WHS mental module on help-seeking behavior when the help of an occupational physician is offered. Besides, the effects on mental health complaints and work functioning are analyzed and discussed in this chapter. In **Chapter 8,** a process evaluation of the WHS mental module is presented, including participants' compliance and perspectives on the WHS mental module as well as occupational physicians' adherence and perspectives on the WHS mental module. This section also addresses suggestions for improvement of the WHS mental module. Finally, **Chapter 9** presents the general discussion. In this chapter, the main research findings are summarized and discussed, methodological considerations are presented, and recommendations for future research and for practical implementation are provided.

REFERENCES

1 Alonso J, Angermeyer MC, Bernert S et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004;21-7.
2 Andrea H, Bültmann U, Beurskens AJ et al. Anxiety and depression in the working population using the HAD Scale - psychometrics, prevalence and relationships with psychosocial work characteristics. *Soc Psychiatry Psychiatr Epidemiol* 2004;39:637-46.
3 Bültmann U, Kant IJ, Kasl SV et al. Lifestyle factors as risk factors for fatigue and psychological distress in the working population: prospective results from the Maastricht Cohort Study. *J Occup Environ Med* 2002;44:116-24.
4 Glozier N. Mental ill health and fitness for work. *Occup Environ Med* 2002;59:714-20.
5 Sanderson K, Andrews G. Common mental disorders in the workforce: recent findings from descriptive and social epidemiology. *Can J Psychiatry* 2006;51:63-75.
6 Watanabe N, Stewart R, Jenkins R et al. The epidemiology of chronic fatigue, physical illness, and symptoms of common mental disorders: a cross-sectional survey from the second British National Survey of Psychiatric Morbidity. *J Psychosom Res* 2008;64:357-62.
7 Bijl RV, Ravelli A, van ZG. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Soc Psychiatry Psychiatr Epidemiol* 1998;33:587-95.
8 Bijl RV, van ZG, Ravelli A et al. The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. *Soc Psychiatry Psychiatr Epidemiol* 1998;33:581-6.
9 Ustun TB, Ayuso-Mateos JL, Chatterji S et al. Global burden of depressive disorders in the year 2000. *Br J Psychiatry* 2004;184:386-92.
10 Bültmann U, Rugulies R, Lund T et al. Depressive symptoms and the risk of long-term sickness absence: a prospective study among 4747 employees in Denmark. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:875-80.
11 Lagerveld SE, Bültmann U, Franche RL et al. Factors associated with work participation and work functioning in depressed workers: a systematic review. *J Occup Rehabil* 2010;20:275-92.
12 Lerner D, Adler DA, Chang H et al. Unemployment, job retention, and productivity loss among employees with depression. *Psychiatr Serv* 2004;55:1371-8.
13 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
14 Wang J, Adair CE, Patten SB. Mental health and related disability among workers: a population-based study. *Am J Ind Med* 2006;49:514-22.
15 World Health Organisation (WHO). World Health Organisation Report 2000. Geneva, Switserland: WHO, 2000.
16 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
17 Aronsson G, Gustafsson K. Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research. *J Occup Environ Med* 2005;47:958-66.
18 Burton WN, Conti DJ, Chen CY et al. The role of health risk factors and disease on worker productivity. *J Occup Environ Med* 1999;41:863-77.
19 Burton WN, Pransky G, Conti DJ et al. The association of medical conditions and presenteeism. *J Occup Environ Med* 2004;46:S38-S45.
20 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
21 Lee YM. Loss of productivity due to depression among Korean employees. *J Occup Health* 2010;52:389-94.
22 McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev* 2009;29:243-59.
23 Waghorn G, Chant D. Work performance among Australians with depression and anxiety disorders: a population level second order analysis. *J Nerv Ment Dis* 2006;194:898-904.
24 Stewart WF, Ricci JA, Chee E et al. Cost of lost productive work time among US workers with depression. *JAMA* 2003;289:3135-44.
25 Suzuki K, Ohida T, Kaneita Y et al. Mental health status, shift work, and occupational accidents among hospital nurses in Japan. *J Occup Health* 2004;46:448-54.
26 Holden RJ, Scanlon MC, Patel NR et al. A human factors framework and study of the effect of nursing workload on patient safety and employee quality of working life. *BMJ Qual Saf* 2011;20:15-24.
27 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.

28 Gershon RR, Stone PW, Zeltser M et al. Organizational climate and nurse health outcomes in the United States: a systematic review. *Ind Health* 2007;45:622-36.
29 Sundin L, Hochwalder J, Bildt C et al. The relationship between different work-related sources of social support and burnout among registered and assistant nurses in Sweden: a questionnaire survey. *Int J Nurs Stud* 2007;44:758-69.
30 Stansfeld S, Candy B. Psychosocial work environment and mental health - meta-analytic review. *Scand J Work Environ Health* 2006;32:443-62.
31 Schultz AB, Edington DW. Employee health and presenteeism: a systematic review. *J Occup Rehabil* 2007;17:547-79.
32 Dewa CS, Lin E. Chronic physical illness, psychiatric disorder and disability in the workplace. *Soc Sci Med* 2000;51:41-50.
33 Motowidlo SJ, Packard JS, Manning MR. Occupational Stress - Its Causes and Consequences for Job-Performance. *J Appl Psychol* 1986;71:618-29.
34 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
35 Wang PS, Beck AL, Berglund P et al. Effects of major depression on moment-in-time work performance. *Am J Psychiatry* 2004;161:1885-91.
36 Borman WC, Motowidlo SJ. Expanding the criterion domain to include elements of contextual performance. In: Schmitt N, orman WC, editors. Personnel Selection in Organizations. San Francisco, CA: Jossey Bass, 1993: 71-98.
37 Koopmans L, Bernaards CM, Hildebrandt VH et al. Conceptual frameworks of individual work performance: a systematic review. *J Occup Environ Med* 2011;53:856-66.
38 Crombie IK, Davies HT. Beyond health outcomes: the advantages of measuring process. *J Eval Clin Pract* 1998;4:31-8.
39 Sonnentag S, Frese M. Performance concepts and performance theory. In: Sonnentag S, editor. Psychological Management of Individual Performance. New York: Wiley, 2002: 3-25.
40 Goetzel RZ, Ozminkowski RJ, Sederer LI et al. The business case for quality mental health services: why employers should care about the mental health and well-being of their employees. *J Occup Environ Med* 2002;44:320-30.
41 Haslam C, Atkinson S, Brown S et al. Perceptions of the impact of depression and anxiety and the medication for these conditions on safety in the workplace. *Occup Environ Med* 2005;62:538-45.
42 Yassi A, Hancock T. Patient safety--worker safety: building a culture of safety to improve healthcare worker and patient well-being. *Healthc Q* 2005;8 Spec No:32-8.
43 World Health Organisation (WHO). International Classification of Functioning, Disability and Health (ICF). Geneva, Switserland: WHO, 2001.
44 Michon HW, ten HM, Kroon H et al. Mental disorders and personality traits as determinants of impaired work functioning. *Psychol Med* 2008;38:1627-37.
45 Birnbaum HG, Kessler RC, Kelley D et al. Employer burden of mild, moderate, and severe major depressive disorder: mental health services utilization and costs, and work performance. *Depress Anxiety* 2010;27:78-89.
46 Hjarsbech PU, Andersen RV, Christensen KB et al. Clinical and non-clinical depressive symptoms and risk of long-term sickness absence among female employees in the Danish eldercare sector. *J Affect Disord* 2011;129:87-93.
47 Abma FI, van der Klink JJ, Terwee CB et al. Evaluation of the measurement properties of self-reported health-related work-functioning instruments among workers with common mental disorders. *Scand J Work Environ Health* 2012;38:5-18.
48 Nieuwenhuijsen K, Franche RL, van Dijk FJ. Work functioning measurement: tools for occupational mental health research. *J Occup Environ Med* 2010;52:778-90.
49 Lam RW, Michalak EE, Yatham LN. A new clinical rating scale for work absence and productivity: validation in patients with major depressive disorder. *BMC Psychiatry* 2009;9:78.
50 Lerner D, Amick Iii BC, Rogers WH et al. The Work Limitations Questionnaire. *Med Care* 2001;39:72-85.
51 Koopman C, Pelletier KR, Murray JF et al. Stanford presenteeism scale: health status and employee productivity. *J Occup Environ Med* 2002;44:14-20.
52 Coles ME, Coleman SL. Barriers to treatment seeking for anxiety disorders: initial data on the role of mental health literacy. *Depress Anxiety* 2010;27:63-71.
53 Dahlberg KM, Waern M, Runeson B. Mental health literacy and attitudes in a Swedish community sample - investigating the role of personal experience of mental health care. *BMC Public Health* 2008;8:8.
54 Putnik K, de JA, Verdonk P. Road to help-seeking among (dedicated) human service professionals with burnout. *Patient Educ Couns* 2011;83:49-54.

55 Lexis MA, Jansen NW, Stevens FC et al. Experience of health complaints and help seeking behavior in employees screened for depressive complaints and risk of future sickness absence. *J Occup Rehabil* 2010;20:537-46.
56 Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv* 2009;60:1051-8.
57 Aromaa E, Tolvanen A, Tuulari J et al. Personal stigma and use of mental health services among people with depression in a general population in Finland. *BMC Psychiatry* 2011;11:52.
58 Sareen J, Jagdeo A, Cox BJ et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. *Psychiatr Serv* 2007;58:357-64.
59 ten Have M, de Graaf R, Ormel J et al. Are attitudes towards mental health help-seeking associated with service use? Results from the European Study of Epidemiology of Mental Disorders. *Soc Psychiatry Psychiatr Epidemiol* 2010;45:153-63.
60 Adams EF, Lee AJ, Pritchard CW et al. What stops us from healing the healers: a survey of help-seeking behaviour, stigmatisation and depression within the medical profession. *Int J Soc Psychiatry* 2010;56:359-70.
61 Siebert DC, Siebert CF. Help seeking among helping professionals: a role identity perspective. *Am J Orthopsychiatry* 2007;77:49-55.
62 Altamura AC, Dell'Osso B, Vismara S et al. May duration of untreated illness influence the long-term course of major depressive disorder? *Eur Psychiatry* 2008;23:92-6.
63 Gordon R. An operational classification of disease prevention. In: Steinberg JA, Silverman MM, editors. Preventing Mental Disorders. Rockville, MD: Department of Health and Human Services, 1987: 20-26.
64 International Labour Organization (ILO). Technical and ethical guidelines for workers' health surveillance. Geneva: ILO, 1998.
65 Koh D, Aw TC. Surveillance in occupational health. *Occup Environ Med* 2003;60:705-10, 633.
66 Netherlands Society of Occupational Medicine (NVAB). Guidance Workers' Health Surveillance [In Dutch: Leidraad Preventief Medisch Onderzoek]. Utrecht, the Netherlands: NVAB, 2005.
67 Plat MJ. Occupational Health Care in High-Demand Jobs: The usefulness of job-specific workers' health surveillance for fire fighters. Amsterdam, The Netherlands: University of Amsterdam, 2011.
68 Sluiter JK. High-demand jobs: age-related diversity in work ability? *Appl Ergon* 2006;37:429-40.
69 Plat MJ, Frings-Dresen MH, Sluiter JK. A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *Int Arch Occup Environ Health* 2011;84:839-57.
70 Winwood PC, Tuckey MR, Peters R et al. Identification and measurement of work-related psychological injury: piloting the psychological injury risk indicator among frontline police. *J Occup Environ Med* 2009;51:1057-65.
71 Levin SM, Herbert R, Moline JM et al. Physical health status of world trade center rescue, recovery workers, volunteers-New York City, July 2002-August. *Morb Mortal Week Report* 2004;53:807-12.
72 Wright KM, Huffman AH, Adler AB et al. Psychological screening program overview. *Mil Med* 2002;167:853-61.
73 Harvey SB, Glozier N, Henderson M et al. Depression and work performance: an ecological study using web-based screening. *Occup Med (Lond)* 2011;61:209-11.
74 Bush K, Kivlahan DR, McDonell MB et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998;158:1789-95.
75 de Beurs E, Zitman FG. The Brief Symptom Inventory (BSI): The reliability and validity of a brief alternative of the SCL-90 [In Dutch: De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90]. *Maandblad GeestelijkeVolksgezondheid* 2006;61:120-41.
76 de Croon EM, Sluiter JK, Frings-Dresen MH. Psychometric properties of the Need for Recovery after work scale: test-retest reliability and sensitivity to detect change. *Occup Environ Med* 2006;63:202-6.
77 Terluin B, van Rhenen W, Schaufeli WB et al. The four-dimensional symptom questionnaire (4DSQ): measuring distress and other mental health problems in a working population. *Work Stress* 2004;18:187-207.
78 van der Klink JJ, van Dijk FJ. Dutch practice guidelines for managing adjustment disorders in occupational and primary health care. *Scand J Work Environ Health* 2003;29:478-87.
79 van der Klink JJ. Guideline: management of mental health problems of workers by occupational physicians [In Dutch: Richtlijn: Handelen van de bedrijfsarts bij werkenden met psychisch problemen]. Utrecht, the Netherlands: NVAB [Netherlands Society of Occupational Medicine], 2007.
80 Nieuwenhuijsen K, Verbeek JH, de Boer AG et al. Validation of performance indicators for rehabilitation of workers with mental health problems. *Med Care* 2005;43:1034-42.

2

The impact of common mental disorders on the work functioning of nurses and allied health professionals: a systematic review

Fania R. Gärtner
Karen Nieuwenhuijsen
Frank J.H. van Dijk
Judith K. Sluiter

International Journal of Nursing Studies 2010;47(8):1047-61.

ABSTRACT

Objective: This study aims to inventory aspects of work functioning of nurses and allied health professionals that are affected by common mental disorders.

Design: A systematic review of psychological and occupational health literature was performed.

Data sources: A sensitive systematic literature search based on index terms and text words was conducted in four electronic databases: PubMed, PsycINFO, Embase, and Cinahl. The literature search was limited to journal articles published between 1998 and 2008, written in English, German, or Dutch.

Review methods: For inclusion, studies had to examine a relationship between common mental disorders and a measure of work functioning in nurses or allied health professionals. No restrictions on study design were handled. Methodological quality was assessed for each study. The data were categorized into themes, for which the strength of evidence was assessed using six levels of evidence.

Results: Sixteen of 2,792 studies met the inclusion criteria, of which 13 had a cross-sectional design, one was a vignette study, and two were narrative reviews. In all studies, the subjects were nurses. The retrieved aspects of sub-optimal work functioning due to common mental disorders were merged into 15 themes. Strong evidence was found for five themes: general errors, medication errors, near misses, patient safety, and patient satisfaction. Moderate evidence was found that common mental disorders are associated with complex motor skills and with general performance; while evidence for an association between common mental disorders and needle stick injuries was inconclusive. Seven themes had only narrative evidence: interpersonal behaviour, energy, focus on goals and responsibility, work speed, avoiding work while on the job, coping with emotions, and motivation.

Conclusion: Common mental disorders were found to be associated with various impairments in work functioning in nurses, these include task-related, intrapersonal and interpersonal aspects of work. In particular, strong evidence was found for an association between common mental disorders and general errors, medication errors, near errors, patient safety, and patient satisfaction. These results provide input for preventive actions to improve both health and work functioning in health care workers.

INTRODUCTION

Work and mental health are in constant interplay. Work can be therapeutic and plays an important role in a person's life; it is a dominant, time consuming, and rewarding activity, and thus a primary source for income, identity, and mental health.[1] However, work and mental health can impact each other in a negative way as well. Unfavourable working conditions can impair the mental health of employees by causing mental health complaints such as stress, mild depression, and anxiety disorders, often referred to as common mental disorders (CMDs).[2] In addition, mental health problems caused by work or other conditions can negatively impact work. CMDs can entail impaired work functioning when being at work, as much as absenteeism from work.[3-5] In the past, absenteeism, its causes and remedies have been studied extensively. More recently, occupational health professionals, work psychologists, and human resource professionals are paying greater attention to the work of employees who are not absent but do suffer from health problems.[6-8] If concrete aspects of work functioning impairments due to CMDs are identified, early detection of employees with CMDs and work functioning problems would be possible, as well as early interventions. This literature review is intended to determine the range of relevant aspects of work functioning that may be impaired by CMDs.

To describe concrete aspects of impaired work functioning, we have chosen to focus on specific occupational groups in the healthcare sector. This review specifically studies nurses, as they are one of the largest occupational groups in this sector. Allied health professionals working in hospitals are also included in this review, as their tasks and work environment are similar to that of nurses in many aspects. By focussing on these occupations too, we hope to find more information on the effect of CMDs on medical actions performed by hospital personal other than doctors. In these occupational groups we can distinguish two risks: first, the increased risk of the onset of CMDs and second, the risk of serious consequences of impaired work functioning for others such as patients. Regarding the first risk, it is noticeable that in the healthcare sector, and in particular for nurses, the incidence of mental health problems is very high.[9] In a study by Wieclaw, the relative risk for depression in nurses was 3.5 (95% CI: 1.3 - 9.6) compared to workers outside the human service professions. The relative risk was highest for nurses compared to other human service workers and also compared to other healthcare workers.[9] A possible explanation for this high prevalence may be high job demands, high emotional demands, a lack of autonomy, and of social support that characterise the work of nurses.[10] These are known risk factors for developing mental health complaints.[11] The second risk, serious consequences of impaired work performance, is also known to be present in nursing.[12] Impaired work functioning can directly affect the health of nurses and their patients. Examples of these effects are needle stick injuries and drug administration errors.[12]

We decided to comprehensively analyse the effect of CMDs on work and chose not to focus solely on the performed actions, or 'output' of the work (e.g., has the nurse given the right pain killer to her patient, at the right time, and with the right dose). We also wanted to take into account all preceding components of this action, the process of work, as it is an essential component of work performance.[13 14] The process of work can be defined as the behaviour exerted to reach work outcomes; thus, what a person is doing at work and how the person is doing it.[14] Regarding the process of pain medication administration, for instance, we may want to know whether the nurse checked the patient's file for necessary information, whether the nurse double checked the medication dose, but also whether the nurse put the patient at ease when administering the medication. By considering work functioning in its entirety, both process and output, a broader and more complete picture is attained of the behaviour of the worker. In the case of impaired functioning, this insight helps to signal the exact impairments and to intervene effectively. Apart from the distinction between work process and work output, we are interested in both the quality and quantity of work performance. In earlier studies, impaired work functioning is most often quantified and expressed in economic terms, such as cost to employers and lost productive time,[15 16] or number of days attending work when unwell.[17] The quality dimension of work has been addressed to a lesser extent, although there are a few examples in which studies focus on errors or accidents.[12 18 19]

Our research question is: what is the effect of CMDs on work functioning of nurses and allied health professionals working in hospitals? To gain insight into the range of relevant aspects of work functioning that are impaired by CMDs, we review information on the output and process of work, as well as quality and quantity aspects.

METHODS

Search strategy

Four electronic databases (PubMed, PsycINFO, Embase, and Cinahl) were systematically searched for peer-reviewed articles published between 1998 and July 2008, and written in English, German, or Dutch. To be representative of the current tasks and work situation of the nurses and allied health professions, as well as of the current opinion on CMDs and their treatment, we limited our search to the past ten years. We used a PICO analysis to develop a systematic search strategy. This resulted in three search groups: 1) patients/ population: nurses and allied health professionals; 2) intervention/ exposure: CMDs; 3) outcomes: work functioning. The first group of search terms represents the patients/ population we studied; i.e., nurses and the 12 allied health professions acknowledged by the Dutch association for allied health professionals. See Table 1 for an overview of the search terms. The second group of search terms stands for the intervention/ exposure, it consisted of terms for CMDs. In

Table 1. The three search categories and its search terms.

PICO search categories	Patients / population: Nurses and allied health professionals	Intervention/ exposure: Common mental disorders	Outcomes: Work functioning
Terms that (or synonyms of which) were used as subject headings and/or text words	Nurses Surgical nurses Nurse anaesthetics Dietician Occupational therapist Physiotherapist Speech therapist Dental hygienist Cesar exercise therapist Mensendieck remedial therapist Orthoptist Podiatrist Radiation therapist Optometrist Skin therapist	Common mental disorder Mental disease Mental illness Psychological disorder Adjustment disorder Anxiety disorder Agoraphobia Obsessive compulsive disorder Panic disorder Post traumatic stress disorders Mood disorder Affective disorder Depression Neurasthenia Sleep disorder Substance-related disorder Alcoholism Burnout Occupational stress Mental fatigue	Mental processes Mental functions Orientation Trust Arousal Attention Wakefulness Memory Psychomotor performance Affect Hardiness Resilience Expressed emotions Perception Thinking Problem solving Decision making adaptation Communication Coping Assisting Interpersonal relations Role functioning Identification Self esteem Self-efficacy Vigilance Responsibility Errors Safety Accident Cognitive deficits Effort Tasks Productivity Dysfunction Job performance Impairments Presenteeism Needle stick injury Work ability

accordance with definitions of Glozier[20] and Watanabe et al.[21] and also based on the prevalence of mental health disorders in the general and working population,[22-24] our definition includes mental health complaints such as burnout, stress, and fatigue, adjustment disorders, anxiety disorders, mood disorders, sleep disorders, and substance-related disorders. The third group embodies outcomes, which are aspects of impaired work functioning. The selection of its search terms was based on an overview of ICF aspects of 'body functions', 'activities', and 'participation' that are expected to be affected by

depression.[25] Furthermore, we added general terms of work functioning and terms specific to nurses work outcomes.

We combined the three groups of search terms with the Boolean operator AND. For each of the three groups of search terms, index terms specific for each database were used and supplemented with free-text words for searching in titles and abstracts. Supplementary to this search strategy, references of included reviews were screened for new relevant titles. Citation tracking using the Science Citation Index was performed on all studies included.

Selection of studies

To be eligible for inclusion, a study had to examine the association between one or more CMDs and a measure of work functioning. Study subjects had to include nurses or allied health professionals. CMDs leading to inclusion were limited to those mentioned in our search strategy. To be as inclusive as possible no restrictions on study design were included. As little previous research is done on our topic, we strive to be as inclusive as possible. Therefore, no restrictions on research design were followed, thus not only primary research data but also narrative reviews were included. Narrative reviews mostly include information that is retrieved from primary studies. However, they also often present data that relies on expert opinions. Nevertheless, this information is empirically derived and appraised by experienced professionals.[26] Therefore, we consider this data as added value for this systematic review." All articles retrieved through our electronic search were evaluated in two steps. In the first step, titles and, where necessary, abstracts of the articles were reviewed. To be eligible for inclusion, the title or abstract either had to exemplify that the study examined CMDs in nurses or allied health professionals or it had to be stated that the study examined an aspect of work functioning in these professions. In this step, first a sample of 10% of the titles and abstracts per database were reviewed independently by two authors (FG, KN). Subsequently, the other 90% were reviewed by one author (FG), and in the case of doubt about possible inclusion the second researcher (KN) was consulted. In the event of difference in the opinion of the two researchers, the article was included to review by full text. In the second step, two researchers (FG, KN) independently reviewed the selected articles by their full text against the inclusion criteria. Disagreements were resolved in a consensus meeting. If no consensus could be reached, a third reviewer (JS) was consulted to aid in the process of consensus-building.

Data extraction

Descriptions of the study population, design, assessment methods, and the results of each included study were summarized in a data extraction table. The results include all aspects of work functioning for which the effect of CMDs were examined in the study. One researcher (FG) performed the data extraction; subsequently a second researcher (KN) checked the extracted data for each study.

Methodological quality assessment

For evaluation of the methodological quality of the included studies, four methodological quality appraisal features were defined related to the sample: the CMD variable, the work function variable, and the statistical methods. These features were assessed using an eight-item quality appraisal format with the two possible scores being '+' and '-'. This list of features was created following a number of guidelines on reporting observational studies (e.g. the STROBE statement) and an example in previous studies.[27-31] Table 2 presents a description of the quality appraisal features that can be adopted for observational research with either cross-sectional or longitudinal design. To asses the validity of recommendations deriving from narrative reviews or vignette studies, three degrees of credibility were handled following the JBI-NOTARI framework.[26] The three degrees of credibility are:

1. unequivocal: evidence beyond reasonable doubt which may include conclusions that are matter of fact, directly reported/ observed and not open to challenge;
2. credible: evidence that is, albeit an opinion, plausible in light of current data and theoretical understandings;
3. unsupported: when neither of the other level descriptors apply and when, most notably, conclusions are not supported by the data.

Two reviewers (FG, KN) independently assessed the methodological quality of the included observational studies and of the narrative reviews and vignette study. Disagreements between the two reviewers were discussed until consensus was reached, if needed with help of a third reviewer (JS).

Data analysis and evidence synthesis

To synthesize the findings of the search, a data-driven thematic analysis of the findings was accomplished.[32] First, the extracted data were categorized into themes sharing common aspects of work functioning via open coding by one author (FG). Second, a definite division was achieved in consensus with all authors.

The data yielded through our search did not allow for meta-analysis. We assessed the strength of evidence of the findings using six defined levels of evidence. These levels depended on three criteria: study design, number of tested relationships per theme, and proportions of significant findings. Regarding study design, longitudinal studies were considered to be of higher evidential value than cross-sectional studies, whereas narrative reviews and vignette studies were considered of minor evidential value. Concerning the number of tested relationships per theme, the more significant findings in the same direction, the higher the level of evidence. Regarding the proportions of significant findings among the tested relationships we decided that for 'evidence', more than 50% of the tested relationships had to lead to significant results in the same direction. The six levels of evidence are as follows:

Table 2. The criteria for methodological quality.

Sample	
Participants	
+	The source of participants* and the method of recruitment is reported.
-	The source of participants* or the method of recruitment is not reported.
Response rate	
+	The response rate is 50% or higher.
-	The response rate is lower than 50% or not reported.
Descriptive data	
+	The main characteristics of the study population (occupation, age, gender) are reported, if applicable per occupational group that is studied.
-	The description of one or more main characteristics of the study population (occupation, age, gender) is missing or not reported per occupational group that is studied.
Disorder or complaints variable	
Description	
+	The common mental disorder variable is defined.
-	The common mental disorder variable is not defined.
Assessment	
+	Details of assessment method** are presented and if applicable categories chosen are listed and explained***.
-	Details of assessment method** are not presented, or, if applicable, categories chosen are not listed and explained***.
Work functioning variables	
Description	
+	The work functioning variable is defined.
-	The work functioning variable is not defined.
Assessment	
+	Details of assessment method** are presented and if applicable categories chosen are listed and explained***.
-	Details of assessment method** are not presented or, if applicable, categories chosen are not listed and explained***.
Statistical methods	
+	The statistics used are described and seem appropriate to achieve the objective of our study.
-	The statistics used are not described and/or seem inappropriate to achieve the objective of our study.

* the group from which the study population was selected
** name and reference of assessment instrument used
*** number of categories and category boundaries when continuous variables were categorized e.g., cut-off points or median values

a) *strong evidence:* significant results in the same direction were found in one or more relationships tested by longitudinal studies, *and* more than 50% of the tested relationships were significant. Or significant results in the same direction were found in at least three relationships tested by cross-sectional studies, *and* more than 50% of the tested relationships were significant.
b) *moderate evidence:* significant results in the same direction were found in two relationships tested by cross-sectional studies, *and* more than 50% of the tested relationships were significant.
c) *limited evidence:* significant results were found in one relationship tested by a cross-sectional study.
d) *expert evidence:* indications from one or more narrative reviews or vignette studies with validity assessed as either “unequivocal” or “credible”.

e) *inconclusive* evidence: significant results were found in at least one relationship tested in a cross-sectional or longitudinal study, however 50% of the tested relationships or less were significant.
f) *inconsistent evidence:* remaining cases, e.g., significant results found in more than one tested relationship pointed in different directions.

RESULTS

Literature search

A total of 3328 titles were identified by the systematic search in four electronic databases. After removing 536 duplicates, 2792 titles and/or abstracts were reviewed for inclusion. This first review step resulted in 63 eligible articles; the full text of these was retrieved. After applying the inclusion criteria to these full text papers, 15 studies were included. Various reasons lead to exclusion of full text articles, e.g., the studied mental health problems were not included in our definition of CMDs, data of different occupational groups were not analysed separately, CMDs were studied as the dependent variable, or the statistical analysis did not give information about the relationship between CMDs and impaired work functioning. The reference check of the included narrative reviews and the citation tracking of the 15 included articles resulted in an additional three potentially eligible studies. Review of the full texts of these three studies led to one further inclusion. Thus, 16 studies met all the criteria for inclusion. Figure 1 shows a flowchart of the inclusion process.

Studies included

As can be seen in Table 3, in which the studies are organized by publication year and then alphabetically, 13 studies applied cross-sectional designs, one was a vignette study presenting cases of impaired work functioning, and two were narrative reviews. In all studies, subjects were nurses working in various wards. One paper studied nurse anaesthetic[33] and another preoperative nurses,[34] in particular. Five studies were conducted in the USA,[33 35-38] three in Canada,[39-41] three in Japan,[42-44] two in Australia,[45 46] and two in Europe (Turkey, The Netherlands).[47 48] From one study the country of origin is unknown.[34] The studies cover six types of CMDs: fatigue (N = 5), stress (N = 4), burnout (N = 5), anxiety (N = 1), depression (N = 1), and addiction illness (N = 1). Most CMDs were assessed by self report. A number of different scales were used to measure mental health. Only one scale, the Maslach Burnout Inventory, was used by several researchers.[35 40 41 44] We have categorized the various aspects of work functioning addressed in the studies into fifteen themes, of which three themes concern errors at work. The fifteen themes are: general errors, medication and equipment errors, needle stick injuries, near misses, patient safety, interpersonal behaviour, patient satisfaction, complex motor skills, general performance, energy, focussing on goals and

responsibility, work speed, avoiding work while on the job, coping with emotions, and motivation. The measures used to assess work functioning varied greatly across the studies. The aspects of work functioning were mainly assessed by self-report questionnaires. Two studies used patient ratings,[38 40] one used self reports with a daily logbook during one month,[45] and one used an observational rating of videotape by an observer and by the nurse subject [37]. Below, the results of the included studies will be described for each of the fifteen themes.

Methodological quality appraisal

Thirteen cross-sectional studies were assessed according to the methodical quality criteria in Table 4. Ten studies fulfilled six or more of the eight criteria,[33 35 38 40-42 44-47] two studies met three to five of the criteria,[37 43] and one study met fewer than three criteria.[34] Only seven studies met the criteria, 'descriptive data' and 'assessment of work functioning', the criteria least frequently met in this review.

The degree of credibility of the two narrative reviews[36 39] was evaluated as "credible", for the vignette study[48] it was "unequivocal".

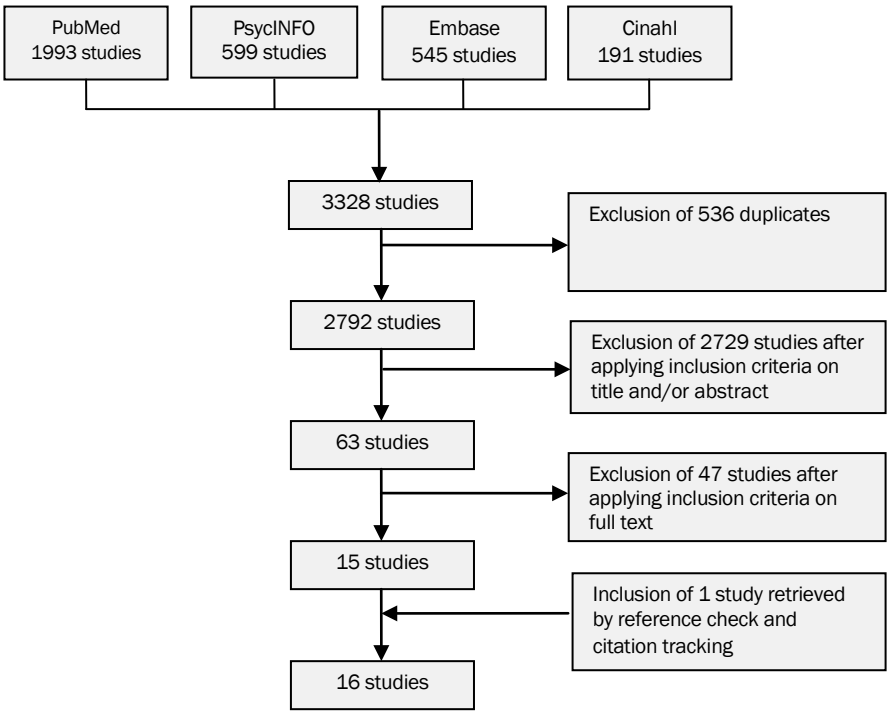


Figure 1. Flow-chart of the inclusion process.

Table 3. Extracted data for the included studies (N = 16).

	Reference	Participants a) occupational group b) sample size (N) c) country of study d) mean age in years (y) e) % females f) ward♦ g) response rate	Study design	Type of CMD or complaints & measurement method a) type of disorder or complaints b) measurement method c) number of respondents with disorder or complaints out of whole sample	Outcome measure & measurement method a) assessed aspect of work functioning b) measurement method c) used measures of association	Results
1.	Dorrian et al. (2008)	a) nurses b) N = 41 c) Australia d) 36 e) 82.9% f) various g) /	cross-sectional	a) stress; physical exhaustion; struggling to remain (STR) awake b) self reported daily record over one month at 6 pm c) not represented on individual level but for a total of 694 shifts: 27% of 694 shifts moderate to high levels of stress, 42% of 694 shifts moderate to high levels of physical exhaustion, 39% of 694 shifts moderate to high levels of mental exhaustion	a) errors and near errors made by oneself or observed b) self report using daily logbooks over one month c) odds ratios (OR) by Multivariate Logistic Regression Analysis	• higher stress rating is a significant predictor of error: OR (95% CI) = 1.48 (1.09-2.02) • higher STR awake rating is a significant predictor of error: OR (95% CI) = 2.40 (1.17-4.91) --- • physical exhaustion is not a significant predictor of error
2.	Halbesleben et al. (2008)	a) nurses b) N = 148 c) USA d) / e) / f) various g) 50%	cross-sectional	a) burnout: • emotional exhaustion • depersonalization b) self report using the Maslach Burnout Inventory (MBI) c) not available	a) patient safety outcomes: • safety grade • safety perception • events reports (12 month recall period) • near-miss frequency reports (12 month recall period) b) self report using the • Agency for Healthcare Research Quality	• emotional exhaustion rating is negatively associated with nurses' perceived patient safety grades (β = -.40**) • depersonalization rating is negatively associated with nurses' perceived patient safety grades (β = -.16*) • emotional exhaustion rating is negatively associated with nurses' safety perception (β = -.84***) • depersonalization rating is negatively associated with nurses' safety perception (β = -.26*) • emotional exhaustion rating is negatively associated with near miss frequency reports (β = -.14*)

Table 3. (continued)

					(AHRQ) • Patient Safety Culture Survey c) β by Multiple Regression Analysis individually for each outcome measure with both burnout variables as predictors	• depersonalisation rating is negatively associated with near miss frequency reports (β = -.36**) --- • emotional exhaustion is not significantly associated with events reported: β = -.02, • depersonalisation is not significantly associated with events reported: β = .01
3.	Kenyon et al. (2007)	a) perioperative nurses b) N = 1013 c) / d) / e) / f) / g) /	cross-sectional	a) fatigue b) not available c) not available	a) performance b) self report c) not available	• 58% reported feeling unsafe while providing patient care • 68% reported they had experienced the effects of sleep deprivation (e.g., being irritable at home and at work, slow response time, family problems, problems sleeping, falling asleep at work); of these 68%, 13% stated they had made patient care mistakes because of fatigue, 38% reported fatigue-related near misses. • patient care-related incidents described by the respondents included: • documentation problems • selecting the wrong implant • not being able to focus while counting, which required repeated partial counts • missed items during patient assessment
4.	Bakker and Heuven (2006)	a) nurses b) N = 108 c) The Netherlands d) 40 e) 81% f) / g) 89%	cross-sectional	a) burnout: • exhaustion • disengagement b) self report using the Oldenburg Burnout Inventory (OLBI) c) not available	a) in-role performance b) self report using a scale of Goodman and Svyantek c) correlation coefficients r	• exhaustion is negatively associated with in-role performance: r = -.26** • disengagement is negatively associated with in-role performance: r = -.22**

Table 3. (continued)

5.	Fogarty and Mckeon (2006)	a) nurses b) N = 176 c) Australia d) / e) 92% f) / g) 63%	cross-sectional	a) individual distress b) self report using the Occupational Positive and Negative Affect Scale c) not available	a) medication violations and medication errors with 12 month recall period b) self report c) correlation coefficients r	• individual distress is positively associated with medication violations: r = .26** • individual distress is positively associated with medication errors: r = .17*
6.	Spence Laschinger and Leiter (2006)	a) nurses b) N = 8.597 (out of 17.965 from 5 countries) c) Canada d) 44 e) 98% f) various g) 59% (regarding total sample of 17.965)	cross-sectional	a) burnout: • emotional exhaustion • depersonalization • personal accomplishment b) self report using the Maslach Burnout Inventory - Human Service Scale (MBI-HSS) c) not available	a) adverse events: frequency of occurrence of 4 kinds of negative patient incidents in their shift: falls; nosocomial infections; medication errors, and patients complaints; with 12 month recall period b) self report c) correlation coefficients r	• emotional exhaustion is positively associated with adverse events: r = .30** • depersonalization is positively associated with adverse events: r = .34** • personal accomplishment is negatively associated with adverse events: r = -.22**
7.	Seki and Yamazaki (2006)	a) nurses b) N = 88 c) Japan d) / e) 99% f) various g) 97.8%	cross-sectional	a) fatigue b) self report on VAS scale before starting work c) not available	a) near miss errors in terms of intravenous medication b) self report after work c) odds ratios(OR) by multiple logistic regression analysis	• lower fatigue before work rating is a significant predictor of near miss errors at day shift: OR (95% CI) = 0.98 (0.95-0.99) --- • fatigue before work rating is not a significant predictor of near miss errors at evening shift: OR (95% CI) = 0.96 (0.92-1.01) • fatigue before work rating is not a significant predictor of near miss errors at night shift: OR (95% CI) = 0.99 (0.94-1.03)
8.	Smith et al. (2006)	a) nurses b) N = 860 c) Japan d) / e) / f) various g) 74%	cross-sectional	a) fatigued after work; stressed from work b) self report c) not available	a) needle stick and sharps injuries (NSI) with 12 month recall period: • Single NSI • Multiple NSI • NSI not reported b) self report	• higher fatigue after work rating is a significant predictor of multiple NSI: OR (95% CI) = 1.87 (1.13-3.13) • higher fatigue after work rating is a significant predictor of NSI not reported: OR (95% CI) = 1.94 (1.03-3.71) ---

Table 3. (continued)

					c) odds ratios(OR) by Logistic Regression Analysis	• fatigue after work rating is not a significant predictor of single NSI: OR (95% CI) = 1.26 (0.83-1.90) • stressed from work rating is not a significant predictor of single NSI: OR (95% CI) = 1.17 (0.79-1.74) • stressed from work rating is not a significant predictor of multiple NSI: OR (95% CI) = 0.73 (0.45-1.18) • stressed from work rating is not a significant predictor of NSI not reported: OR (95% CI) = 0.97 (0.53-1.75)
9.	Vahey et al. (2004)	a) nurses b) N = 820 c) USA d) 34.6 e) 92.6% f) various g) 86%	cross-sectional	a) burnout • emotional exhaustion • depersonalization • personal accomplishment b) self report using Maslach Burnout Inventory (MBI) c) not available	a) patient satisfaction b) patient interview using the La Monica-Oberst Patient Satisfaction Scale (LOPSS) (patients: N = 621 with response rate of 86%) c) linear regression analysis, before and after adjusting for patients' sex, age, race, risk factors, and illness severity; burnout rates were aggregated to unit level	• lower emotional exhaustion rating is a significant predictor of patient satisfaction: OR (95% CI) = 0.51 (0.30-0.87) • higher personal accomplishment rating is a significant predictor of patient satisfaction: OR (95% CI) = 2.37 (1.37-4.12) --- • depersonalisation rating is not a significant predictor of patient satisfaction: OR (95% CI) = 1.21 (0.76-1.91)
10.	Pilette (2005)	a) nurses b) / c) USA d) / e) / f) / g) /	narrative review	a) depression b) not applicable c) not applicable	a) presenteeism b) not applicable c) not applicable	• failed collaboration • low morale • errors • signs of depression in the workplace • increased irritability with patients, families, and co-workers • fatigue, daydreaming, tardiness • increased on-the-job accidents and

Table 3. (continued)

						errors • frequently missed deadlines • working more slowly • loss of interest in activity • increased social isolation from co-workers • according to items of a questionnaire by Jones: The employee reports, he or she is or is not • able to finish work tasks • able to focus on goals and responsibilities • energetic enough to complete work • repeating jobs or tasks • working more slowly than usual • experiencing problems with supervisor or peers regarding the amount of work he or she is doing • having accidents or making errors at work
11.	Suzuki et al. (2005)	a) nurses b) N = 4279 (out of 4407 nurses in total with both sexes) c) Japan d) 30.3 e) 100% f) Various g) 94% (regarding total sample of 4407)	cross-sectional	a) excessive daytime sleepiness b) self report using Pittsburgh Sleep Quality Index c) not available	a) occupational accidents with 12 month recall period: • drug administration error • incorrect operation of medical equipment • needle stick injury b) self report univariate analysis and multiple logistic repression analysis adjusted for age	• excessive daytime sleepiness is a significant predictor for incorrect operation of medical equipment: OR (95% CI) = 1.27 (1.09-1.49) --- • excessive daytime sleepiness is not a significant predictor for drug administration errors: OR = 1.19 (0.97-1.29) • excessive daytime sleepiness is not a significant predictor for needle stick injuries: OR (95% CI) = 1.13 (0.98-1.31)

Table 3. (continued)

12.	Adlersberg and MacKinnon (2004)	a) nurses
		b) /
		c) Canada
		d) /
		e) /
		f) /
		g) /
			narrative review	a) addiction illness
				b) not applicable
				c) not applicable
					a) behaviour at the job and job performance
					b) not applicable
					c) not applicable
						• unusual behavioural patterns
						  • offering to work more overtime than usual
						  • frequently being absent from work, late for work or use vague excuses to take sick time
						  • taking long or frequent breaks
						  • mood fluctuations
						  • sleeping on the job
						  • wanting to be alone
						  • reporting a chaotic home life
						  • feeling picked on at work
						  • over-reacting emotionally to situations
						• job performance
						  • more errors and omissions than usual
						  • doing just enough to get by at work
						  • increase in the number of complaints from other nurses, doctors, and/or patients. when challenged may offer implausible excuses for own behaviour or become defensive

13.	Karadeniz and Cakamakci (2002)	a) nurses b) N = 43 c) Turkey d) 24.44 e) / f) / g) /	vignette study	a) possible causes for medication errors b) assessed by one open question c) not available	a) medication error b) described in five scenarios c) percentages of causes answered	• 33% of respondents report tiredness and exhaustion of nurse as main cause for medication errors
14.	Smith et al. (2001)	a) nurses b) N = 26 c) USA d) 27.8 e) / f) intensive care	cross-sectional	a) state anxiety b) self report by state/trait anxiety inventory (STAI) c) not available	a) complex motor skills b) assessment of videotaped ICU ‘endotracheal (ET) suctioning’ during a work sample and in	• high state anxiety predicts ICU ET suctioning performance self-rated by nurse: p = .04, performance score went down by 10% (β coefficients not reported by authors)

Table 3. (continued)

		g) unit 57%			the ICU rated by oneself and by supervisor based on a 10 step protocol c) linear regression analysis	
15.	Kendrick (2000)	a) nurses anaesthetists b) N = 15 (out of total sample of 81, other 66 were students) c) USA d) 40.4 e) 53% f) level 1 trauma centre g) /	cross-sectional	a) occupational strain b) self report using the Personal Strain Questionnaire (PSQ) c) not available	a) communication style b) self report using the Interpersonal Communication Inventory (ICI) c) multiple regression analysis	• no significant results --- • occupational strain is no significant predictor of communication style
16.	Leiter et al. (1998)	a) nurses b) N = 711 (out of 1268 nurses from total sample of 3312) c) Canada d) / e) 96.7% f) various g) 83% (of whole sample)	cross-sectional	a) burnout: • exhaustion • cynicism • professional efficacy b) self reports using the Maslach Burnout Inventory – general survey (MBI-GS) c) not available	a) patient satisfaction rating of nurses b) patient rating using the Patient Judgement of Hospital Quality questionnaire (assessed by 931 patients, response rate 58.%) c) correlation coefficients r	• exhaustion rating is negatively associated with patient satisfaction rating of nurses: r = -.73* • cynicism rating is negatively associated with patient satisfaction rating of nurses: r = -.53* --- • professional efficacy rating is not significantly associated with patient satisfaction rating of nurses: r = .06

♦ more than three different wards are called 'various'
* $p < .05$
** $p < .01$

Aspects of suboptimal work functioning

General errors

The relationship between CMDs and general errors was the subject of three cross-sectional studies.[35 41 45] Five of the eight studied relations (62.5%) showed significant results in the same direction. Results show that CMDs are associated with the occurrence of more errors. ORs (95% CI) of these findings were 1.5 (1.09-2.02) and 2.4 (1.17-4.91), correlation coefficients varied between 0.22 and 0.34 in the described direction. In sum, there is strong evidence for the relationship between CMDs and general errors. Additionally, expert evidence stemming from two narrative and one vignette study supports the relationship between CMDs and errors. In the narrative reviews,[36 39] impact on errors by CMDs is described with the following terms: 'increased on-the-job accidents and errors', 'having accidents or making errors at work', and 'more errors and omissions than usual'. In the vignette study it is found that 33% of the respondents report nurses' tiredness and exhaustion as a primary cause for medication errors.[48]

Descriptive data add information to these findings. One study yields that of 68% of respondents who reported that they had experienced effects of sleep deprivation, 13% stated that they had made patient care mistakes due to fatigue.[34] Furthermore, the descriptive data adds information about the type of patient care-related incidents related to sleep deprivation, which are 'documentation problems', 'selecting the wrong implant', 'not being able to focus while counting, which required repeated partial counts', and 'missed items during patient assessment'.[34]

Medication and equipment errors

Two studies focused on the relationship between CMDs and medication and drug administration errors.[44 46] Outcome variables were defined as 'medication violations', 'medication errors', 'incorrect operation of medical equipment', and 'drug administration error'. In total, three of the four tested relationships (75%) between CMDs and medication errors were significant and positive. Correlation coefficients were 0.17 and 0.26, the OR (95% CI) was 1.3 (1.09-1.29). In conclusion, strong evidence was found for the relationship between CMD and medication errors.

Needle Stick Injuries

The association between CMDs and needle stick injuries was the subject of two studies,[43 44] in which the following outcome measures were used: 'single needle stick injury', 'multiple needle stick injuries', 'not reported needle stick injury', and 'general needle stick injury'. In total, two of seven associations between CMDs and needle stick injury were significantly positive (29%). Both reported ORs (95% CI) were 1.9 with confidence intervals of (1.13-3.13) and (1.03-3.71), none were in the other directions. This implies inconclusive evidence.

Table 4. Methodological quality of cross-sectional studies (N = 13).

Reference	Sample			Common mental disorder variable		Work functioning variables		Statistical methods
	Partici-pants	Response rate	Descrip-tive data	Descrip-tion	Assess-ment	Descrip-tion	Assess-ment	
Dorrian et al. (2008)	+	-	+	+	-	+	+	+
Halbesleben et al. (2008)	+	+	-	+	+	+	+	+
Kenyon et al. (2007)	+	-	-	-	-	-	-	-
Bakker and Heuven, (2006)	+	+	+	+	+	+	+	+
Fogarty and Mckeon (2006)	+	+	+	+	+	+	-	+
Spence Laschinger and Leiter (2006)	+	+	+	+	+	+	+	+
Seki and Yamazaki (2006)	+	+	-	+	+	+	-	+
Smith et al. (2006)	+	+	-	-	-	-	-	+
Vahey et al. (2004)	+	+	+	+	+	+	+	+
Suzuki et al. (2005)	+	+	+	+	+	+	-	+
Smith et al. (2001)	-	+	-	+	+	+	-	+
Kendrick (2000)	+	-	+	+	+	+	+	+
Leiter et al. (1998)	+	+	-	+	+	+	+	+

Near misses

Two studies focused on the relationship between CMD and near miss errors.[35 42] Outcome measures were 'near miss frequency reports', 'near miss errors with intravenous medication at day shift', 'near miss errors with intravenous medication at evening shift', and 'near miss errors with intravenous medication at night shift'. Three out of the five studied relationships between CMDs and near miss errors had significant results (60%), indicating a negative association between CMDs and near errors. β varied from -0.14 to -0.36, the reported OR (95% CI) was 0.98 (0.95-0.99). In conclusion, strong evidence was found for the relationship between the prevalence of CMD and less near miss errors. In contrast to this strong evidence, descriptive data adds that 38% out of 68% respondents who report experiences with effects of sleep-deprivation reported fatigue-related near misses.[34]

Patient safety

Two studies examined the relationship between CMDs and patient safety outcomes.[34 35] All four studied associations were significantly negative (100%), thus more CMDs were related to less patient safety. β ranged from -0.16 to -0.84.[35] There is strong evidence for this relationship. Descriptive data adds that in the context of sleep deprivation, 58% of the respondents reported feeling unsafe while providing patient care.[34]

Interpersonal behaviour

One study examined interpersonal behaviour outcomes,[33] with the outcome measure 'communication style'. However, the one tested association was not significant. Furthermore, a negative relationship between CMDs and interpersonal behaviour was mentioned in two narrative reviews.[36 39] The narrative reviews give the following examples of impaired interpersonal behaviour due to CMDs: 'irritability with patients, families, and co-workers', 'increased isolation from co-workers', 'failed collaboration',[36] 'wanting to be alone', and 'more complaints from co-workers and/or patients'.[39] This leads to expert evidence for the relationship between CMDs and interpersonal behaviour.

Patient satisfaction

The effect of CMDs on patient satisfaction was tested in two studies.[38 40] In total, four of the six tested associations were significant (67%). The results all imply a negative association between CMDs and various measures of patient satisfaction. ORs (95% CI) were 0.5 (0.30-0.87) and 2.4 (1.37-4.12), Spearman correlation coefficients varied from -0.53 to -0.73. In conclusion, there is strong evidence that CMDs are associated with lower patient satisfaction.

Complex motor skills

The impact of CMDs on the technical performance of a complex motor skill was tested in one study,[37] with a self and supervisor rating of ICU endotracheal suctioning as the outcome

measures. For both tested associations a significant effect was found (100%), suggesting that a higher level of CMDs was associated with lower performance rating (the β coefficient was not reported by the authors). Thus, moderate evidence for this relationship was found.

General performance

One study examined the association between CMDs and general work outcomes.[47] The outcome measures was 'in-role performance'. In total, two of the two tested associations were significant (100%), these results imply a negative relationship between CMDs and performance, with correlation coefficients of -0.22 and -0.26. These findings lead to the conclusion of moderate evidence for the relationship between CMDs and overall performance.

Energy

The theme energy in relation to CMDs is discussed in one narrative review.[36] It is described as 'not being energetic enough to complete work'. Expert evidence suggests a negative relationship between CMDs and energy.

Focussing on goals and responsibility

According to two narrative reviews,[36 39] CMDs influence the ability to focus on goals and sense of responsibility. This was demonstrated in descriptions of examples: 'not being able to finish tasks', 'frequently miss deadlines', and 'discussions about amount of work', and 'doing just enough to get by at work'. In conclusion, the negative relationship between CMDs and focusing on goals and being responsible is supported by expert evidence.

Work speed

One narrative review describes slower work speed as an effect of CMDs.[36] Examples discussed were 'working more slowly', 'daydreaming', and 'tardiness'. Therefore, expert evidence suggests a negative relationship between CMDs and work speed.

Avoiding work while on the job

Avoiding work while on the job is discussed in one narrative review.[39] Examples are 'being late', 'taking longer breaks', and 'sleeping on the job'. Thus, expert evidence was given for the positive relationship between CMDs and being absent during work.

Coping with emotions

Decreased coping with emotions is discussed in one narrative review.[39] This theme is described by terms such as 'mood fluctuations', 'overreact emotionally to situations', 'becoming defensive', and 'feeling picked on at work'. Thus, expert evidence is given for the negative association between CMDs and coping with emotions.

Motivation

Motivation in relation to CMDs is discussed in one narrative review.[36] This theme is illustrated by 'low morale' and 'loss of interest in activities'. This expert evidence suggests a negative relationship between CMDs and motivation.

In Table 5 an overview of the level of evidence for each theme is given.

DISCUSSION

A systematic search in four databases and a reference check identified 16 papers on the effect of CMDs on work functioning in nurses. The data yielded was synthesized into 15 themes of impaired work functioning, for which the level of evidence was assessed. Strong evidence was found for five themes: general errors, medication errors, near misses, patient safety, and patient satisfaction. Moderate evidence was found for two themes: complex motor skills and general performance. Inconclusive evidence was found for the theme needle stick injuries. Seven themes received solely narrative evidence: interpersonal behaviour, energy, focus on goals and responsibility, work speed, avoiding work while on the job, coping with emotions, and motivation. For all themes, results indicate a positive association between CMDs and impaired work functioning, except for near misses.

The variety of outcomes clearly point to the heterogeneous nature of the effect of CMDs on work functioning. A number of aspects of work functioning are represented, such as task related results (errors, near misses, motor skills, work speed), results that describe intrapersonal aspects of work (avoiding work while on the job, energy, motivation, coping with emotions, focusing on goals and responsibility), and results that indicate interpersonal aspects of work. To our knowledge this is the first review that aimed to systematically inventory impairments in work functioning due to CMDs in nurses and allied health professionals. In addition to its innovative topic, the strength of this review is the extensive search among four databases, as well as our broad approach of work functioning. To gain insight into the wide range of impaired work functioning aspects, we focused on process and output as well as on quantitative and qualitative work aspects. The following themes yielded from our data synthesis are examples of process aspects of work functioning: near misses, interpersonal behaviour, motor skills, energy, focussing on goals and responsibility, work speed, avoiding work while on the job, coping with emotions, and motivation. The remaining themes are examples of output aspects of work functioning: patient safety, patient satisfaction, and general performance. While all results found were of qualitative nature, the error themes make distinctions between process and output and between qualitative and quantitative aspects of work difficult. Errors are process-related and have a qualitative character. However, the way the data is presented in the studies does not give insight into the

Table 5. Overview of level of evidence per theme.

Theme		Level of evidence					
		Strong	Mode-rate	Limited	Incon-clusive	Narra-tive	Incon-sistent
1.	General errors	x				x	
2.	Medication and equipment errors	x					
3.	Needle stick injuries				x		
4.	Near misses	x					
5.	Patient safety	x					
6.	Interpersonal behaviour					x	
7.	Patient satisfaction	x					
8.	Complex motor skills		x				
9.	General performance		x				
10.	Energy					x	
11.	Focussing on goals and responsibility					x	
12.	Work speed					x	
13.	Avoiding work while on the job					x	
14.	Coping with emotions					x	
15.	Motivation					x	

process of work. In the data found, errors are expressed in quantitative terms, as the number of errors was measured. Evidence of the qualitative character of errors, such as differentiation of error types, severity, and point in time, is limited. For the assessment of delivered quality of care and for the development of high quality care, more qualitative information would be valuable, e.g., descriptions of what exactly was done (incorrectly) and when.[49]

A remarkable finding here was that fewer symptoms of CMD are associated with more near misses. Near errors give the impression of impairments in functioning, however the researchers of the studies emphasised that the assessment concerned measures of the detection and report of near misses.[35 42] Thus, near misses here can be regarded as signs of reflection and attention to the subjects' own work functioning.

Methodological considerations

The included studies had various methodological limitations, of which four are distinct. The first concerns the design of the included studies. There is an obvious lack of longitudinal studies, as no cohort studies were found. As all yielded data were association coefficients, conclusions about causality between CMDs and impairments in work functioning should be drawn with caution. Yet it has to be noted that it is questionable whether longitudinal designs are the most adequate designs when studying this topic. Work functioning is directly related to mental state in time, thus measurements of CMDs and work functioning with a time gap are not necessarily suitable to determine the causal relationship. A second methodological restriction regards the assessment of work functioning. Of the 14 studies for which the assessment method was reported, 11 studies used self report measures. Probable bias due to self reports can over-exaggerate associations, since subjects with CMDs may rate their own functioning less favourably than healthy persons, as has been demonstrated in depression.[49] A third point of attention regards the sometimes ambiguous definitions of work impairments.

In two studies error was not directly linked to the CMDs of the nurse reporting these errors.[41 45] Rather, errors were defined as "error and near errors made or observed by oneself" and "negative patient incidents on their shifts". This might lead to a bias in the number of errors. A fourth restriction points to missing information in most studies about the distribution of the CMDs reported. A lack of sufficient variation in CMDs in the study samples may be one reason for the absence of significant association coefficients in some of the studies. This may lead to solely 'inconclusive evidence' of an association for some themes.

The most important methodological consideration of this systematic review regards the use of levels of evidence to summarize our findings. This process is useful when the data do not allow a meta-analysis; however, it is also controversial because of its arbitrary decisions. It should be used with caution, as the sizes of effects found are not included as they are in meta-analysis.[50] Even though this form of determining strength of evidence is not free from critique, the process is transparent, as the exact meanings of the levels are described. This enables readers to draw their own conclusions. With regard to our choice to consider the percentage of significant associations in the assignment of the level of evidence, we note that underpowered studies may interfere with the strength given by larger studies that do find significant results. Therefore, we emphasise that the label 'inconclusive evidence' does not imply the absence of a relationship. Rather, it highlights the fact that more insignificant results than significant were found.

Furthermore, we want to note that the used framework for the quality assessment of the narrative data, the JBI-NOTARI, lacks explicit assessment criteria. Thus, the quality evaluation of this data is made on a subjective judgement. To our knowledge no more detailed guidelines for the quality evaluation of narrative data exist. Therefore, we want to point out the need for the development of more detailed quality assessment criteria for narrative data.

Finally, it is noteworthy that all results were derived from studies with nurses as subjects. No studies were found which focus on allied health professionals. Therefore, the conclusions we draw are based on nurse data only.

Practical relevance

This review can be of practical relevance for employees and employers as well as for further intervention development. Firstly, the gathered knowledge could result in more awareness in nurses and their supervisors of impairments in work functioning. This awareness might be sufficient to decrease the risk of adverse events, e.g., because employees pay more attention on tasks or because managers accommodate tasks of employees to their individual impairments. Secondly, the results can be of use in the development of detection and intervention tools, e.g., detection instruments. Mental health problems of employees are often not detected and treated in time, partly because symptoms in some cases are not yet severe enough to enable a diagnosis. But even early symptoms of CMDs can lead to impairments in work functioning.[51] By screening for work impairments due to CMDs, the chance of signalling

of problems early is increased. In this way, both health problems and work functioning problems in nurses and allied health professionals can be reduced.

Implications for further research

More information is needed about less-studied or unstudied aspects of impaired work functioning due to CMDs, in particular in allied health professionals. Further research should focus on the role of symptom severity and its relationship to work functioning impairments. Questions such as 'how severe do mental health problems need to be to affect work functioning?' and 'what is the effect of symptom reduction on existing work functioning impairments?' remain unanswered. Moreover, the results of this review can be used for the development of a screening instrument that focuses on the whole range of impairments in work functioning in the studied occupations.

Conclusion

In conclusion, we can state that CMDs are associated with various impairments in the work functioning of nurses. Overall, these impairments concern task related functioning aspects, intrapersonal aspects and interpersonal aspects of work. In particular, strong evidence is found for associations between CMDs and general errors, medication errors, near errors, patient safety, and patient satisfaction. Results point to a positive relationship between CMDs and work functioning impairments, except for near misses. These results provide input for preventive actions to improve both health and work functioning in health care workers.

REFERENCES

1 Tsang H, Chiu IY. Development and validation of the workshop behavior checklist: a scale for assessing work performance of people with severe mental illness. *Int J Soc Psychiatry* 2000;46:110-21.
2 Wieclaw J, Agerbo E, Mortensen PB et al. Psychosocial working conditions and the risk of depression and anxiety disorders in the Danish workforce. *BMC Public Health* 2008;8:280.
3 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
4 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
5 Bonde JP. Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. *Occup Environ Med* 2008;65:438-45.
6 Aronsson G, Gustafsson K. Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research. *J Occup Environ Med* 2005;47:958-66.
7 Burton WN, Conti DJ, Chen CY et al. The role of health risk factors and disease on worker productivity. *J Occup Environ Med* 1999;41:863-77.
8 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
9 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.
10 Gershon RR, Stone PW, Zeltser M et al. Organizational climate and nurse health outcomes in the United States: a systematic review. *Ind Health* 2007;45:622-36.
11 Stansfeld S, Candy B. Psychosocial work environment and mental health - meta-analytic review. *Scand J Work Environ Health* 2006;32:443-62.
12 Suzuki K, Ohida T, Kaneita Y et al. Mental health status, shift work, and occupational accidents among hospital nurses in Japan. *J Occup Health* 2004;46:448-54.
13 Crombie IK, Davies HT. Beyond health outcomes: the advantages of measuring process. *J Eval Clin Pract* 1998;4:31-8.
14 Sonnentag S, Frese M. Performance Concepts and Performance Theory. In: Sonnentag S, editor. Psychological Management of Individual Performance. New York: Wiley, 2002.
15 Schultz AB, Edington DW. Employee health and presenteeism: a systematic review. *J Occup Rehabil* 2007;17:547-79.
16 Stewart WF, Ricci JA, Chee E et al. Cost of lost productive work time among US workers with depression. *JAMA* 2003;289:3135-44.
17 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
18 Chambers R, George V, McNeill A et al. Health at work in the general practice. *Br J Gen Pract* 1998;48:1501-4.
19 Haslam C, Atkinson S, Brown S et al. Perceptions of the impact of depression and anxiety and the medication for these conditions on safety in the workplace. *Occup Environ Med* 2005;62:538-45.
20 Glozier N. Mental ill health and fitness for work. *Occup Environ Med* 2002;59:714-20.
21 Watanabe N, Stewart R, Jenkins R et al. The epidemiology of chronic fatigue, physical illness, and symptoms of common mental disorders: a cross-sectional survey from the second British National Survey of Psychiatric Morbidity. *J Psychosom Res* 2008;64:357-62.
22 Bijl RV, Ravelli A, van ZG. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Soc Psychiatry Psychiatr Epidemiol* 1998;33:587-95.
23 Bijl RV, van ZG, Ravelli A et al. The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. *Soc Psychiatry Psychiatr Epidemiol* 1998;33:581-6.
24 Sanderson K, Andrews G. Common mental disorders in the workforce: recent findings from descriptive and social epidemiology. *Can J Psychiatry* 2006;51:63-75.
25 Cieza A, Chatterji S, Andersen C et al. ICF Core Sets for depression. *J Rehabil Med* 2004;128-34.
26 Pearson A. JBI-NOTARI Users Manual for version 4.0. Australia: The Joanna Briggs Institute, 2004.
27 Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377-84.
28 Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666-76.
29 Vandenbroucke JP, Von EE, Altman DG et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* 2007;18:805-35.

30 Von Elm E, Altman DG, Egger M et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344-9.
31 Wind H, Gouttebarge V, Kuijer PP et al. Assessment of functional capacity of the musculoskeletal system in the context of work, daily living, and sport: a systematic review. *J Occup Rehabil* 2005;15:253-72.
32 Dixon-Woods M, Agarwal S, Jones D et al. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy* 2005;10:45-53.
33 Kendrick P. Comparing the effects of stress and relationship style on student and practicing nurse anesthetists. *AANA J* 2000;68:115-22.
34 Kenyon TA, Gluesing RE, White KY et al. On call: alert or unsafe? A report of the AORN on-call electronic task force. *AORN J* 2007;86:630-9.
35 Halbesleben JR, Wakefield BJ, Wakefield DS et al. Nurse burnout and patient safety outcomes: nurse safety perception versus reporting behavior. *West J Nurs Res* 2008;30:560-77.
36 Pilette PC. Presenteeism in nursing: a clear and present danger to productivity. *J Nurs Adm* 2005;35:300-3.
37 Smith AM, Ortiguera SA, Laskowski ER et al. A preliminary analysis of psychophysiological variables and nursing performance in situations of increasing criticality. *Mayo Clin Proc* 2001;76:275-84.
38 Vahey DC, Aiken LH, Sloane DM et al. Nurse burnout and patient satisfaction. *Med Care* 2004;42:II57-II66.
39 Adlersberg M, MacKinnon J. Registered nurses and substance misuse or abuse: RNABC's role. *Nursing BC* 2004;36:13-5.
40 Leiter MPH. The correspondence of patient satisfaction and nurse burnout. *Soc Sci Med* 1998;47:Nov.
41 Spence Laschinger HK, Leiter MP. The impact of nursing work environments on patient safety outcomes: the mediating role of burnout/engagement. *J Nurs Adm* 2006;36:259-67.
42 Seki Y, Yamazaki Y. Effects of working conditions on intravenous medication errors in a Japanese hospital. *J Nurs Manag* 2006;14:128-39.
43 Smith DR, Mihashi M, Adachi Y et al. Epidemiology of needlestick and sharps injuries among nurses in a Japanese teaching hospital. *J Hosp Infect* 2006;64:44-9.
44 Suzuki K, Ohida T, Kaneita Y et al. Daytime sleepiness, sleep habits and occupational accidents among hospital nurses. *J Adv Nurs* 2005;52:445-53.
45 Dorrian J, Tolley C, Lamond N et al. Sleep and errors in a group of Australian hospital nurses at work and during the commute. *Appl Ergon* 2008;39:605-13.
46 Fogarty GJ, McKeon CM. Patient safety during medication administration: the influence of organizational and individual variables on unsafe work practices and medication errors. *Ergonomics* 2006;49:444-56.
47 Bakker AB, Heuven E. Emotional Dissonance, Burnout, and In-Role Performance Among Nurses and Police Officers. *Int J Stress Manage* 2006;13:423-40.
48 Karadeniz G, Cakmakci A. Nurses' perceptions of medication errors. *Int J Clin Pharmacol Res* 2002;22:111-6.
49 Wang PS, Beck AL, Berglund P et al. Effects of major depression on moment-in-time work performance. *Am J Psychiatry* 2004;161:1885-91.
50 de Vet HC, van Tulder MW, Bouter LM. Levels of evidence: intellectual aid or absolute judgement? *J Clin Epidemiol* 2003;56:917-8.
51 Yassi A, Hancock T. Patient safety–worker safety: building a culture of safety to improve healthcare worker and patient well-being. *Healthc Q* 2005;8 Spec No:32-8.

3

Impaired work functioning due to common mental disorders in nurses and allied health professionals: the Nurses Work Functioning Questionnaire

Fania R. Gärtner
Karen Nieuwenhuijsen
Frank J.H. van Dijk
Judith K. Sluiter

International Archives of Occupational and Environmental Health 2012;85(2):125-38.

ABSTRACT

Purpose: Common mental disorders (CMD) negatively affect work functioning. In the health service sector not only the prevalence of CMDs is high, but work functioning problems are associated with a risk of serious consequences for patients and healthcare providers. If work functioning problems due to CMDs are detected early, timely help can be provided. Therefore, the aim of this study is to develop a detection questionnaire for impaired work functioning due to CMDs in nurses and allied health professionals working in hospitals.

Methods: First, an item pool was developed by a systematic literature study and five focus group interviews with employees and experts. To evaluate the content validity, additional interviews were held. Second, a cross-sectional assessment of the item pool in 314 nurses and allied health professionals was used for item selection and for identification and corroboration of subscales by explorative and confirmatory factor analysis.

Results: The study results in the Nurses Work Functioning Questionnaire (NWFQ), a 50-item self-report questionnaire consisting of seven subscales: *Cognitive aspects of task execution*, *Impaired decision making*, *Causing incidents at work*, *Avoidance behavior*, *Conflicts and irritations with colleagues*, *Impaired contact with patients and their family*, *Lack of energy and motivation*. The questionnaire has a proven high content validity. All subscales have good or acceptable internal consistency.

Conclusion: The Nurses Work Functioning Questionnaire gives insight into precise and concrete aspects of impaired work functioning of nurses and allied health professionals. The scores can be used as a starting point for purposeful interventions.

INTRODUCTION

Mental health complaints such as stress, mild depression, and anxiety disorders, often referred to as common mental disorders (CMDs), can lead to impairments in work performance.[1-5] These impairments result not only in lower productivity; but in certain occupations they can have serious consequences as well, e.g., in the work of nurses and allied health professionals. In these professions, consequences of impaired work functioning can affect the health of the caregiver as well their patients. Examples of these deleterious effects include medication errors, needle stick injuries, near errors, and decreased patient satisfaction.[6] These consequences are even more noteworthy given the high incidence of CMDs in this occupational group. The relative risk of depression is highest for nurses, RR = 3.5 (95% CI: 1.3 – 9.6), as compared to other human service workers and other healthcare workers.[7] A forecasted increase in the shortage of nursing personnel may result in an even higher risk for the development of CMDs, due to increased job demands and patient load or lack of supervisor support.[8-11]

Both the high prevalence of CMDs and the high risk of serious adverse events in these occupations call for action. If we know the exact aspects of work functioning that are impaired, we can purposefully intervene in a proactive manner. In the short run, knowledge of impairments could result in increased awareness on the part of the employee, the supervisors, and the managers, which might be a starting point for discussion and personal support. Also, help-seeking behavior might be stimulated by the insight into impaired work functioning. Finally, detection of problems in work functioning due to CMDs can guide in developing purposeful interventions to improve work functioning and contribute to solutions for underlying mental health problems. For this purpose, sound measuring instruments can be helpful.

Examples of measuring instruments such as questionnaires for assessing impairments in work functioning do exist: the Work Limitation Questionnaire (WLQ),[12] the Stanford Presenteeism Scale (SPS),[13] [14] and the Endicott Work Productivity Scale (EWPS).[15] However, the detection ability of these scales has not been studied.[16] We assume that mild CMDs can also result in impaired work functioning, even though the worker might not always be aware of the presence of mental health problems and their consequences. Many of the existing work functioning scales, e.g., the WLQ and the SPS, explicitly refer to health problems in their items. However, these questionnaires are less suitable for detecting new cases of workers with impaired work functioning due to mental disorders.

Furthermore, existing instruments were developed for the work context in general, rather than for a specific occupational group.[17] An advantage of focusing on specific occupations is that items in a measuring instrument can refer more directly to the actual work practice and to concrete experiences of the employees. This approach enables the detection of specific aspects of work functioning that are impaired and thus enables subsequent concrete

interventions. Therefore, we aim to develop a questionnaire for the early detection of impaired work functioning due to CMDs in nurses and allied health professionals. Our research questions are as follows:

1. Which self-report questionnaire items can be formulated to detect CMD-associated impairments in the work functioning of nurses and allied health professionals and how is the content validity of these questionnaire items evaluated by the target population?
2. Which subscales can be distinguished from a pool of formulated items and which items can be selected for definite use in the questionnaire?

Since a universally accepted definition of work functioning is missing, different concepts and perspectives were discussed by the authors to guide the development process of the detection instrument. Based on these various conceptions, for the purposes of our study, we consider work functioning as a comprehensive concept, encompassing a wide range of aspects measurable by self-reports. We include aspects of the work process and work outcome,[18] as well as aspects of task execution and of organizational functioning, such as behavior within the team and towards the environment of the work organization.[19 20] Additionally, the extra effort to complete work tasks is included where appropriate.[21] Furthermore, in the present study, rather than expressing impairments of work functioning solely in terms of quantity, qualitative aspects of work functioning will be addressed as well.[22-24] Following this description, we assume work functioning to be a multidimensional construct; therefore, no prior limit was set on the number of subscales and items the instrument should contain. Yet, we strive to develop a self-report questionnaire based on the classical test theory assumptions.

In the following, the methods and results of the two research questions will be described separately as part 1 and part 2.

METHODS

Methods part 1: Development of the item pool

Design

In order to develop a sound questionnaire with high content validity, a protocol based on recommendations by Haynes et al. [25] and by Terwee et al. [26] was followed. The development of the item pool comprised of three phases: the preparation phase, the item generation phase and the revision phase, is described in detail below. Figure 1 presents an overview of the study design with the methods and results for each step.

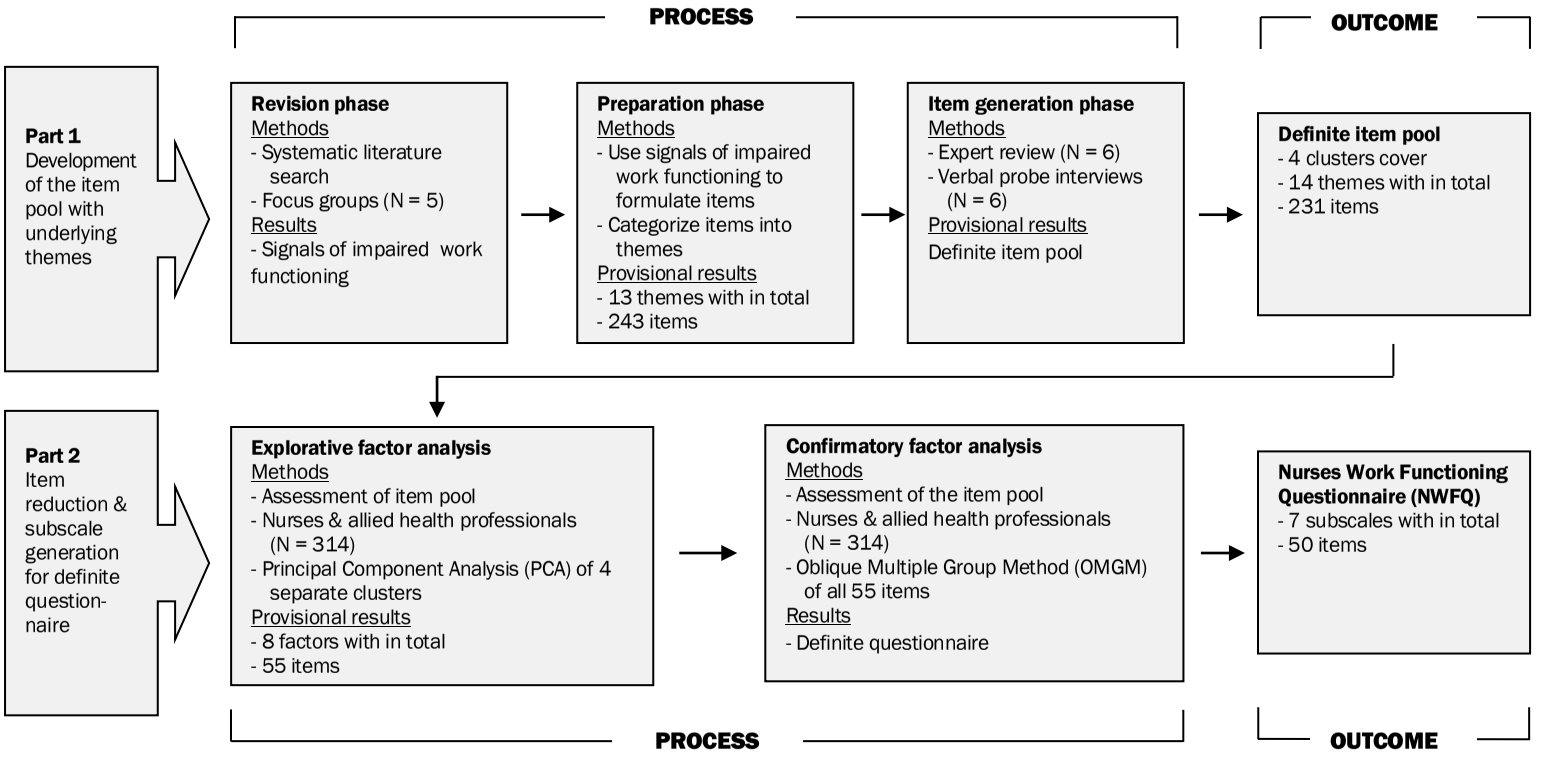


Figure 1. Overview of the study design and the results of each step.

Preparation phase

Procedure of the preparation phase

In the first phase, we conducted two systematic literature searches in four databases: PubMed, PsycINFO, Embase, and Cinahl. We aimed to inventory all literature about effects of CMDs on work functioning in general (first search) and nurses and allied health professionals in particular (second search).[6] Subsequently, five focus group interviews were held. Following a multiple category design,[27] three focus groups were held with nurses and allied health professional and two with experts on work functioning in the health sector.

The focus group interviews with a duration of 2 hours were conducted by two researchers (FG & KN) who alternately moderated or observed. The group interviews were structured by three cases, which were presented to the participants. The cases, written in the second person, described, respectively, an employee with fatigue and stress, depression and anxiety, and alcohol abuse. Participants were asked to reflect on aspects of the work that might be affected by the mental health complaints described. By working with these cases, participants of the employee focus groups were not forced to disclose whether mentioned examples were derived from own experiences or from the behavior of colleagues. In the beginning of each focus group, the discussion was explorative in nature. Later on, aspects of impaired work functioning derived from our literature review were validated and supplemented with illustrative examples. The moderator ensured that for each aspect of impaired work functioning mentioned, the different occupations and specialties present gave concrete examples. The moderator explicitly asked for differences in experiences between the various occupational groups present. Also, the moderator asked to clarify any ambiguities in the examples of participants.

Each focus group discussion was audio taped. The Medical Ethics Committee of the Academic Medical Center Amsterdam decided that approval of the research protocol by the committee was not required.

Subjects of the preparation phase

Focus group members were recruited from one academic medical center using a purposive sampling procedure, with variation in wards and occupations as a major criterion. Nurses and allied health professionals for the three employee focus groups were invited via head nurses. For the selection of participants in the focus groups, we asked for a mix between healthy participants and participants with current or past mental health complaints. We assumed that every employee can deliver input on the research question either based on own experiences with mental health complaints in the presence or past or otherwise based on observations in co-workers with mental health complaints. Employees from the same ward were assigned to different focus groups. Information was collected about the participants' history of mental health complaints. Of the 19 participants, 16 had experienced a difficult period in life with

effects on their mental health in the past and three currently experienced problems. Nine participants had (mild) mental health complaints in the past and one currently had.

Participants for the expert focus groups, such as senior nurses and occupational physicians, were personally invited. Informed consent was obtained from each participant, and all participants were compensated with a 25 Euro voucher for their 2-h participation.

Analysis of the preparation phase

Audiotapes of the focus groups were transcribed verbatim. The analysis of the focus group interviews followed a purpose-driven approach, aiming to distinguish as many different signals of impaired work functioning as possible and to organize all signals into themes.[27] First, each interview was open coded. In this inductive step, all examples of impairments in the work functioning were indexed. During the coding procedure, we aimed to be as inclusive as possible. Therefore, in case of inconsistencies between codes, no exclusion or broadening of codes was performed but inconsistent codes were preserved. Second, codes were refined and reduced within a process of re-reading and constant comparison.[28] Third, the obtained codes were categorized into themes covering related aspects of work functioning. One researcher (FG) performed the coding of the data; subsequently, a second researcher (KN) checked the coded data of each interview. For the analysis of the literature review, see Gärtner et al.[6]

Textbox. Cases used for the focus group discussion.

Case 1:

Try to imagine yourself in the following situation:

Due to conflicts at home you have not been feeling well the past weeks. You have much less energy than usual and after a long day at work you feel too exhausted to do your everyday activities and to relax. This morning you arrive at work feeling stressed already, today will be a very busy day again. Just the idea of all the work you have to do makes you tired.

What difficulties do you expect to face during this workday?

Case 2:

Try to imagine yourself in the following situation:

Since a few months you have not been feeling very well. In the last few weeks you have been feeling especially bad. You feel depressed, there is nothing you want to do or what excites you. The only thing you feel like doing is to stay in your bed all day long. At work you sometimes feel anxious without any reason; you can't tell where the anxiety comes from, the feelings just comes over you. In the past weeks you have had more and more difficulties to accomplish your tasks at work.

Can you describe how your working day goes in these circumstances?

Case 3:

Try to imagine yourself in the following situation:

You have a nice team you work with, with many different people and you get along with each other very well. Since a while you have noticed that one of your colleagues behaves differently. Regularly, you have the feeling she smells of alcohol.

What has changed in the behavior of your colleague?

Item generation phase

Procedure of the item generation phase

In the second phase, items were formulated based on the results of the literature search and focus groups. For each theme that resulted from the preparation phase, sufficient items for possible subscales were formulated (minimum of seven). Each item had to refer to a clear, concrete single action or behavior. To connect with the actual behavior and perception of nurses and allied health professionals, item formulation had to reflect expressions from focus group participants as much as possible. Where possible, items had to be applicable for the different tasks and jargons of the various occupations and specialties as well. A four-week timeframe was chosen for all items. Response formats were chosen according to the content of the associated themes with a minimum of five and maximum of seven categories.[29] Response scales utilized were the following: Likert-type scales (from 0 = *totally disagree* to 6 = *totally agree*), an adjectival scale asking for difficulty (from 0 = *no difficulty* to 6 = *great difficulty*), relative frequency categories (from 0 = *almost never* to 6 = *almost always*), or absolute frequency categories from (0 = *not once* to 6 = *on average more than 1x per day*).

Analysis of the item generation phase

The results of the focus groups together with the information derived from the literature reviews were synthesized into themes and all signals of impaired work functioning were translated into items. These were discussed several times by all of the authors, which resulted in the first pool of items. In this phase, we adhered to the principle of being as inclusive as possible.[26]

Revision phase

Procedure of the revision phase

As part of the revision phase, the first pool of items was submitted for an expert check. Six experts (head nurses and occupational health professionals) were asked to identify items that were unclear or irrelevant. They were asked to rate the relevance of each theme and the completeness of the questionnaire as a whole on a 5-point Likert scale ranging from 1 = *not at all relevant/complete* to 5 = *highly relevant/complete*. On item level, the relevance was rated on a 2-point scale (*yes*, *no*). In addition, participants were invited to suggest supplementary themes and items.

Subsequently, verbal probe interviews were conducted with six nurses and allied health professionals who reviewed the individual items in a 1-hour interview.[30] Participants were asked to identify any item that was unclearly formulated, difficult to respond to, or not applicable to all nursing wards and allied health professions. Additionally, the preference for response formats was discussed.

Subjects of the revision phase

For the expert checks, six key persons (head nurses and occupational health professionals) were invited. For the verbal probe interviews, six nurses and allied health professionals were invited personally. The sampling in this phase was again purposive and we aimed to have as many different professions represented, e.g., also (head) nurses form anesthetic and surgical nursing wards and allied health professionals. The experts, nurses, and allied health professionals invited, partly already participated in the focus group interviews and partly were newly recruited.

Analysis of the revision phase

Possible changes in the item pool resulting from the expert checks and verbal probe interviews were proposed by one researcher (FG) and discussed by the research team until consensus was reached. Items and response categories that were reworded where when possible checked in subsequent interviews. Expert comments on missing signals of impaired work functioning led to the formulation of additional items. In order to draw conclusions on the content validity, the quantitative results about the relevance and clarity of themes and items were summarized by frequencies of the given answers.

Methods part 2: item reduction and subscale generation

Procedure part 2

The second part of our study has a cross-sectional design. Respondents were contacted by e-mail and asked to fill out an electronic version of the item pool, which took approximately 45 min for completion on a computer. It was possible to log out half way through the survey and to continue after logging in again later on. However, the questionnaire had to be fully completed within 3 days. It was not possible to skip questions. Two reminders to complete the questionnaire were sent by e-mail. For each completed questionnaire, we donated 2.50 Euro to a charity that the respondents could select from among three options.

Subjects part 2

A random sample of 1,200 nurses and allied health professionals in one Dutch academic medical center was taken, as we expected a response rate of 25% and strived to recruit 300 respondents. This sample was stratified by age, gender, and occupation.

Information was collected about the participant's gender, age, and the history of their mental health complaints. Mental health status was measured using two questionnaires. First, the *General Health Questionnaire* (GHQ-12) was used, a 12-item self-report questionnaire developed to detect CMDs in the general population.[31] Following earlier studies in the working populations, a cut-off point of ≥ 4 was applied to identify individuals reporting sufficient psychological distress to be classified as probable cases of minor psychiatric disorder.[32]

Second, the 16 item distress subscale of the *Four-Dimensional Symptoms Questionnaire* (4DSQ) was used.[33 34] For case identification, a cut-off point of ≥ 11 was applied.[35]

Analysis part 2

A first reduction in items was based on the variation in answers. In the case of minimal variation (≥95% of answers given in one response category), exclusion of the item was discussed in the research team.[29] Further reduction in items and determination of the underlying factors were based on explorative factor analysis with an orthogonal rotation approach, using principal component analysis (PCA) and varimax rotation.[36 37] To determine the optimum number of factors, we considered Catell's screetest.[38] Kaiser's criterion (retain factors with Eigenvalue > 1),[39] and parallel analysis, following the criterion that the PCA Eigenvalue of our dataset had to exceed the mean Eigenvalue of 100 random datasets with the same number of items and sample size.[40] In cases where these methods led to different numbers of components, we preferred the most interpretable component structure, with the least number of components.

Subsequently, we performed a sequence of PCA Varimax rotations and the analysis of internal consistency (using Cronbach's alpha coefficients), to give meaning to the selected factors, to distribute items to the factors, and to further reduce the number of items.[36 41] Items were assigned to a factor if their factor loading was 0.40 or greater.[36] In case of cross-loadings, they were assigned to the factor with highest factor loading. The selection of items forming the definite subscale was based on the following considerations:

1. The content of the items: selected items should clearly represent the subconstruct with as many different facets as possible.
2. Factor loading: items with higher factor loadings were preferred.
3. Cronbach's alpha: items with highest contribution to the scale's overall alpha were proposed for selection.

The analyses were repeated after each deletion of items until the unidimensional structure of each subscale was stable without further improvement in the alpha coefficient. A Cronbach's alpha of at least 0.70 was regarded sufficient and above 0.80 as good.[29 42]

Since the item pool was too large (231 items) to analyze in one PCA, we analyzed four clusters of themes that are related to each other from a theoretical point of view. This division is in line with existing models of job performance.[20] Our first cluster, "cognitive aspects of work functioning", corresponds with the idea of task performance. The second cluster, "causing incidents", corresponds with counterproductive behavior, although we do not regard causing incidents as voluntary, which is part of the definition of counterproductive behavior. Our third cluster, "interpersonal behavior", and fourth cluster, "energy and motivation" are in accordance with organizational performance and the extra effort needed to perform the work, respectively. See Table 2 for the allocation of themes to the clusters.

Finally, to test whether the selected subscale structure remained stable, a confirmatory factor analysis with all remaining items from all clusters was carried out, using the Oblique Multiple Group Method.[43 44] Based on the highest item test correlations for each item on each subscale, it can be determined for which subscale the individual items have the best fit. Possible incorrect assignments of items to subtests were corrected in this step.

All statistical analyses were performed using SPSS version 16.0, except for the Parallel Analysis, which was conducted using Monte Carlo PCA for Parallel Analysis.[45]

RESULTS

Results part 1: development of the item pool

The literature reviews together with the five focus groups initially yielded 13 themes of impaired work functioning with underlying items. The themes resulting from the systematic literature review and the focus groups overlapped to a large extent. However, the focus group data provided more detailed themes on task execution and comprehensive examples of behavior for all themes.

The focus groups further revealed that reflection on and insight in ones own behavior and insufficiencies at work is difficult for employees with mental health complaints. As a result, we formulated the proposed items as concrete examples of behavior, whenever possible. Also, where appropriate, we chose for response formats with frequency categories, to help respondents rely on memories of past behavior rather than on an interpretation of it.

The 13 themes with underlying items were evaluated in the revision phase by the expert checks. For the relevance element of content validity, themes were evaluated by means of a 5-point scale question. For nine themes, 100% of the scores were given in categories either 4 or 5. For the other four themes, at least a two-thirds majority of answers were in categories 4 or 5 (see Table 1). Thus, all themes were assessed from relevant to highly relevant. The second element of content validity, comprehensive representation of the targeted construct, was evaluated by means of a 5-point scale question. Two of the six given answers were in category 4, and the other four answers were in category 5. Therefore, the representativeness of the item pool was assessed as highly comprehensive. One new theme derived from the expert checks was “execute work less independently”, under which relevant items from other themes or newly formulated items were organized.

On the item level, the revision phase led to the addition of eight new items and the deletion of 20 original items, mainly due to overlap or ambiguity. Further comments in this phase led to re-wordings of items. One example of rephrasing was the change of the term “errors” into “incidents”, as this term more explicitly indicates the involuntary nature of these unintended actions.

After the revision phase, the item pool consisted of 14 themes with a total of 231 items. These themes were grouped into four clusters. See Table 1 for the themes and a description of the items. Figure 1 presents an overview of the results for each step of this study.

Results part 2: item reduction and subscale generation

The socio-demographic characteristics and the mental health complaints of the sample with 314 subjects are presented in Table 2. The sample is representative of the occupational groups, working in the academic medical center where our sample was recruited.

Item reduction by explorative factor analysis

As expected, all 231 items had a highly skewed distribution of answers. First, 19 items were deleted because of too little variance in answers. The data of all four clusters were suitable for the PCA. However, the PCA for the second cluster (causing incidents) had to be performed without the data of the allied health professionals, as too many "not applicable to my job" answers were given in this group, leading to too many missing values. The Kaiser-Meyer-Olkin values for the four clusters were 0.73, 0.72, 0.80, and 0.90, respectively; all exceeding the recommended value of 0.60.[46 47] Bartlett's test of sphericity was significant in all cases (with $p < 0.0001$).[48] Table 3 presents an overview of PCA results and a description of the content of the items included per selected factor. In the supplemented files, we present the rotated component matrix with the factor loadings for each cluster.

The PCA of the first cluster was performed with 82 items, of which 19 remained. Based on the scree-plot and the interpretability of the factors, a three-factor solution was chosen. It accounted for 32% of the explained variance. The following subscales were identified: "cognitive aspects of task execution", "withdrawing from responsibilities", and "impaired decision making".

The PCA of the second cluster was performed with 41 items, of which 15 remained. An interpretable one-factor solution was chosen based on the scree-plot, explaining 23% of the total variance. The identified subscale was "causing incidents at work".

For the third cluster, out of 61 items, 19 remained. The scree-plot of the PCA pointed to four factors, which were highly interpretable. It accounted for 36% of the overall variance. Subscale one is "avoiding contact with colleagues" and two is "conflicts and irritations with colleagues". Subscale three and four are "impaired contact with patients and their family"; because of their overlap in underlying content, they were combined.

In the PCA of the fourth cluster, with 28 items of which six remained, we chose the one-factor solution, based on the scree-plot and the good interpretability. It explains 35% of the variance. This subscale is called "lack of energy and motivation".

Table 1. The generated themes grouped by four clusters to be used in the PCA.

Cluster (number of items in item pool)	**Themes** (number of items in item pool)	**Behavior described by the items of the theme**	**Relevancy** [frequencies of given answers (five response categories: 1='not at all relevant' to 5='extremely relevant') (experts: N = 6)]
Task execution (91)	Attention and concentration problems (14)	Concentration, focusing on tasks, being able to multitask, working precise and accurate	4 (N = 2) 5 (N = 4)
	Losing the overview (17)	Being able to prioritize in tasks, thinking ahead, having the overview of all work tasks, being able to anticipate to unforeseen situations	4 (N = 3) 5 (N = 3)
	Impaired memory (8)	Forgetting important things at work, control behavior due to doubts about which tasks are done yet	3 (N = 1) 4 (N = 1) 5 (N = 4)
	Impaired decision making (12)	Not being able to make decisions themselves, hand decision-making tasks over to co-workers, lack of trust in the quality of own decisions at work	4 (N = 2) 5 (N = 4)
	Working less efficient (11)	Not being able to complete work in time, needing more time to finish tasks, being less productive at work	3 (N = 1) 4 (N = 3) 5 (N = 2)
	Limiting one's tasks (23)	Avoiding busy shifts or difficult patients, giving up special tasks like supervision tasks, the absence of extra role behavior like helping co-workers	2 (N = 1) 3 (N = 1) 4 (N = 3) 5 (N = 1)
	Execute work less independently (6)	Being in need for help from co-workers to ask questions, to finish work, or to avert incidents	*added after revision phase*
Causing incidents (46)	Causing incidents (46)	Causing incidents in general as well as different sorts of incidents, i.e., medication administration, documentation, communication	4 (N = 2) 5 (N = 4)
Inter-personal behavior (65)	Contact with patients and their relatives (26)	Speaking in an inappropriate tone to patients or relatives, being impatient, having lack of empathy, avoiding difficult or emotional situations with patients, not being able to prevent conflicts with patients or relatives	2 (N = 1) 4 (N = 4) 5 (N = 1)
	Aggressive behavior (11)	Rough treatment of patients and co-workers, blaming patients for unsuccessful care	4 (N = 3) 5 (N = 3)
	Impaired contact with colleagues and supervisors (19)	Avoidance of contact with co-workers, becoming irritated and angry about organisational issues, conflicts with co-workers	4 (N = 1) 5 (N = 5)
	Avoid work and colleagues while on the job (9)	Avoidance of talks, contact and collaboration with co-workers and supervisors, withdrawal from common rooms to be alone	4 (N = 5) 5 (N = 1)
Experience of work and emotions at work (29)	Experience work to be more demanding (8)	Having trouble managing the work load, more energy needed to execute work, feeling the need for extra days off	4 (N = 2) 5 (N = 3)
	Emotions (21)	Having feelings of losing control at work, being anxious, being short tempered, becoming emotional, being unsure about the own skills, being unmotivated	4 (N = 2) 5 (N = 4)

Table 2. Participant characteristics (N = 314).

Demographic characteristics		
Gender (N (%))		
female	257	(81.2)
male	57	(18.2)
Age in years (mean (SD))	44.5	(12.0)
Marital status (N (%))		
married/ living together with a partner	227	(72.3)
being in a relationship	21	(6.7)
single	54	(17.2)
divorced	11	(3.5)
widow/ widower	1	(0.3)
Ethnical background (N (%))		
Dutch	261	(83.1)
immigrant first generation	35	(11.1)
immigrant second generation	18	(5.7)
Occupation (N (%))		
nurse	220	(70.1)
surgical nurse	23	(7.3)
anaesthetic nurse	13	(4.1)
allied health professional	58	(18.5)
Working experience in years (mean (SD))	20.8	(12.2)
Kind of contract (N (%))		
permanent position	301	(95.9)
fixed-term contract	9	(2.9)
temporary employment	4	(1.3)
Work hours per week (mean (SD))	30	(6.3)
Mental health complaints	83	(26)

For each cluster, a final PCA was performed with the selected items. For all clusters, the selected number of factors was corroborated.

Confirmatory factor analysis

A confirmatory factor analysis was performed on all 55 items of the eight subscales that remained after the PCA. Overall, the distribution of items into the subscales was confirmed. Some items have high scores on a subscale with which their own subscale is highly correlated. We regard these correlations as acceptable, as long as the score on its own subscale is higher or close. The results of the Oblique Multiple Group Method led to combining of two subscales, “withdrawing from responsibilities” and “avoiding contact with colleagues”, into a new subscale named “avoidance behavior”. Also, a total of four items were replaced and five were removed. In the supplemented files, we present the rotated component matrix with the factor loadings for each cluster. At the end of this study, a questionnaire with seven subscales and a total of 50 items was derived (Table 4). The internal consistency is good in four subscales (0.81-0.94) and acceptable in three subscales (0.70-0.78).

The first subscale was “cognitive aspects of task execution and general incidents”, covering eleven items on working efficiently, alertly, accurately, independently, keeping track of the tasks, and causing incidents in general. The second subscale is “impaired decision making”. This subscale encompasses three items regarding the ability to make important and

Table 3. Results of the PCA for all four clusters.

Subscale	Content of items included for this subscale	# of items	N*	Cronbach's α	# of items with factor-loading < 0.4	# of items with cross-loading > 0.3	% explained variance**
Cluster 1							
Cognitive aspects of task execution	Working efficiently, wakeful, accurate, independently, and not forgetting things at work; keeping track of one's tasks	7	311	0.92	0	0	69.0
Withdrawing from responsibilities	Avoiding specific tasks, decision making, and responsibilities	8	264	0.61	0	0	33.6
Impaired decision making	The ability to make important and quick decisions in stressful situations and knowing solutions for unexpected problems	4	306	0.75	1	0	62.5
Cluster 2							
Causing incidents	Incidents in general	11	197	0.84	1	-	37.4
Causing general incidents***	Different types of incidents, i.e., medication administration, documentation and interpretation	4	326	0.74	0	-	56.4
Cluster 3							
Avoiding contact with colleagues	Avoidance of talks and cooperation with and the presence of co-workers and supervisors	4	309	0.87	0	0	72.0
Conflicts and irritations with colleagues	Feelings of anger and irritation regarding co-workers; conflicts and tensions in the team	7	311	0.82	0	2	42.6
Impaired contact with patients and their family	Lack of time, empathy for and patience with patients and their relatives	8	223	0.81	0	2	43.7
Cluster 4							
Lack of energy and motivation	Lack of motivation, experiencing work as more demanding	6	307	0.82	0	-	58.1

* number of respondents who answered all items
** percentage of variance explained by the first factor in each subscale
*** this subscale is a selection of items from the subscale 'causing incidents' which are applicable to allied health professionals

Table 4. Psychometric properties of the definite seven subscales.

Subscale	# of items	N*	Cronbach's α	Theoretical range of sum score	Range of sum score in sample (median)
Cognitive aspects of task execution and general incidents	11	308	0.94	0-100	0-82 (5)
Impaired decision making	3	310	0.88	0-100	0-100 (0)
Causing incidents at work**	8	176	0.78	0-100	0-40 (4)
Avoidance behavior	8	294	0.70	0-100	0-81 (0)
Conflicts and irritations with colleagues	7	311	0.77	0-100	0-61 (4)
Impaired contact with patients and their family	8	223	0.81	0-100	0-42 (4)
Lack of energy and motivation	5	307	0.81	0-100	0-73 (7)

* number of respondents who answered all items, this N is used for Cronbach's α and the range of the sum score in the sample

** data of nurses only is analyzed

quick decisions in stressful situations. The third subscale was "causing incidents at work", consisting of the eight items covering different types of incidents: medication administration, documentation, and interpretation. This scale was not suitable for the allied health professionals, as too many of them answered "not applicable to my job" on more specific incidents items. The fourth subscale was "avoidance behavior", which encompassed eight items about avoiding particular tasks and responsibilities as well as avoiding contact and cooperation with co-workers. The fifth subscale was "conflicts and irritations with colleagues", its seven items described feelings of anger and irritation regarding co-workers and conflicts and tensions in the team. The sixth subscale was "impaired contact with patients and their family", that included eight items about lack of time, patience, and empathy for patients and their family. The seventh subscale was "lack of energy and motivation". Its five items refer to lack of motivation and experiencing work as more demanding. Table 4 presents the psychometric properties of the seven subscales. We present the definite questionnaire in the "Appendix".

DISCUSSION

Aim of this study was to develop a job-specific detection questionnaire for impaired work functioning due to CMDs in nurses and allied health professionals. In the first part of this study, various signals of impaired work functioning due to CMDs were identified, using literature research and focus group interviews and later translated into items. These signals covered 14 themes of work functioning impairments and described concrete behavior or actions of the work of nurses and allied health professionals. In the second part, seven clear and interpretable factors were distinguished by factor analysis, grouping 50 items of the original 231 items in the item pool. Four of the seven subscales have good alpha's (above 0.80), three have acceptable alpha's (above 0.70). Based on the evaluations from the expert

check and verbal probe interviews, we conclude that the content validity of our instrument is high. The newly developed questionnaire is called the Nurses Work Functioning Questionnaire (NWFQ).

The development of the questionnaire followed a clear step-by-step procedure, planned in advance. In the development process, we used literature as well as qualitative data presenting knowledge and experiences of employees and experts as input sources. Furthermore, in the quality assessment of possible items and the choice of definite items and subscales, both expert opinions and statistical analyses were used. In conclusion, the procedure employed exemplifies the requirements for the development of a scientific questionnaire that is relevant for practice.[25 26]

The focus group interviews were applied as one step in this development study. Using a purposive sampling strategy, the focus group data include experiences from diverse nursing specialisms and experts' professions. Therefore, we assume that the focus group results are applicable to the whole spectrum of the work of nurses and allied health professionals. This comprehensive approach is an important aspect of quality in qualitative research methods.

Unlike existing work functioning scales, the NWFQ aims to be job-specific. It comprises aspects of work functioning that are not, or are to a lesser extent, included in generic work functioning instruments. One specific aspect is "causing incidents". In healthcare service, incidents can have serious consequences for the health of patients as well as for the health of the workers. Therefore, detecting a high risk of incidents is indispensable when assessing (impaired) work functioning in nurses and allied health professionals. A second aspect, which exemplifies the value of job-specific scales, regards interpersonal behavior. Most generic questionnaires do not include this aspect of work functioning as a separate subscale, such as the Endicott Work Productivity Scale (EWPS)[15] or use only a few items without the distinction between co-workers and other persons, e.g., the Work Limitations Questionnaire (WLQ).[12] The quality of communication with patients and their family forms a crucial element of the NWFQ, as this work aspect is essential in the health service sector. Not only does the job-specific approach lead to more concrete examples of behavior in the items itself, it also leads to a better coverage of the most relevant aspects of the work. Therefore, the job-specific approach used here is of additional value to similar measurement instruments that approach work functioning more generally.

Based on insights from the focus groups that reflection on one's own behavior is sometimes insufficient when suffering from mental health complaints, we aimed to formulate items that present behavior as concrete as possible. However, as the items also had to be broad enough to be applicable to the different nursing wards, some items give room for broader interpretation. For example, the item on assessing which (nursing) care a patient needs (item 30) can relate, e.g., to giving the right decubitus prophylaxis, delivering the right medication, or choosing correct patients' transport.

Implementation of the questionnaire should await the results of further research on its construct validity and reproducibility. Also, to draw conclusions about the detection ability of the NWFQ, results on the discriminative validity are necessary.

The multidimensionality of the instrument and the nature of the items allow for more accurate assessment of the nature of impairments in work functioning. High scores provide a starting point for purposeful interventions. Depending on the specific aspects and severity of impairments, interventions can be tailored. Interventions can be of small scale, such as paying more attention to the specific (impaired) work aspects or by a temporarily adjustment of tasks. Interventions can also be of larger scope, guided by professional counselors such as psychologists or occupational health physicians. Future research should focus on 1) the implementation of various interventions using the NWFQ and 2) the effectiveness of those interventions.

Conclusion

The Nurses Work Functioning Questionnaire (NWFQ), a 50-item multidimensional measure of impaired work functioning in nurses and allied health professionals due to CMDs, was developed. Its seven subscales, with high content validity and good internal consistency, cover the full range of impaired work functioning of nurses and allied health professionals with CMDs. The individual subscale scores give insight into the precise aspects of impaired work functioning, allowing for tailoring of interventions for individual needs. Therefore, the NWFQ demonstrates both breadth and depth of measurement, while allowing for self-administration within a reasonable amount of time.

REFERENCES

1 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
2 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
3 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
4 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
5 McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev* 2009;29:243-59.
6 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
7 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.
8 Gershon RR, Stone PW, Zeltser M et al. Organizational climate and nurse health outcomes in the United States: a systematic review. *Ind Health* 2007;45:622-36.
9 Stansfeld S, Candy B. Psychosocial work environment and mental health - meta-analytic review. *Scand J Work Environ Health* 2006;32:443-62.
10 Sundin L, Hochwalder J, Bildt C et al. The relationship between different work-related sources of social support and burnout among registered and assistant nurses in Sweden: a questionnaire survey. *Int J Nurs Stud* 2007;44:758-69.
11 Virtanen M, Pentti J, Vahtera J et al. Overcrowding in hospital wards as a predictor of antidepressant treatment among hospital staff. *Am J Psychiatry* 2008;165:1482-6.
12 Lerner D, Amick Iii BC, Rogers WH et al. The Work Limitations Questionnaire. *Med Care* 2001;39:72-85.
13 Koopman C, Pelletier KR, Murray JF et al. Stanford presenteeism scale: health status and employee productivity. *J Occup Environ Med* 2002;44:14-20.
14 Turpin RS, Ozminkowski RJ, Sharda CE et al. Reliability and validity of the Stanford Presenteeism Scale. *J Occup Environ Med* 2004;46:1123-33.
15 Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull* 1997;33:13-6.
16 Nieuwenhuijsen K, Franche RL, van Dijk FJ. Work functioning measurement: tools for occupational mental health research. *J Occup Environ Med* 2010;52:778-90.
17 Sanderson K, Tilse E, Nicholson J et al. Which presenteeism measures are more sensitive to depression and anxiety? *J Affect Disord* 2007;101:65-74.
18 Sonnentag S, Frese M. Performance concepts and performance theory. In: Sonnentag S, editor. Psychological Management of Individual Performance. New York: Wiley, 2002: 3-25.
19 Motowidlo SJ, Van Scotter JR. Evidence that task performance should be distinguished from contextual performance. *J Appl Psychol* 1994;79:475-80.
20 Viswevaran C, Ones DS. Perspectives on Models of Job Performance. *Int J Select Assess* 2000;8:216-26.
21 Dewa CS, Lin E. Chronic physical illness, psychiatric disorder and disability in the workplace. *Soc Sci Med* 2000;51:41-50.
22 Haslam C, Atkinson S, Brown S et al. Perceptions of the impact of depression and anxiety and the medication for these conditions on safety in the workplace. *Occup Environ Med* 2005;62:538-45.
23 Suzuki K, Ohida T, Kaneita Y et al. Mental health status, shift work, and occupational accidents among hospital nurses in Japan. *J Occup Health* 2004;46:448-54.
24 Yassi A, Hancock T. Patient safety--worker safety: building a culture of safety to improve healthcare worker and patient well-being. *Healthc Q* 2005;8 Spec No:32-8.
25 Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assessment* 1995;7:238-47.
26 Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
27 Krueger RA, Casey MA. Focus groups: a practical guide for applied research. Thousand Oaks, CA: Sage Publications, 2000.
28 Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000;320:114-6.
29 Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. fourth ed. Oxford: Oxford University Press, 2008.

30 Willis GB. Cognitive interviewing in practice: think-aloud, verbal probing and other techniques. Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks, CA: Sage Publications, 2005: 42-63.
31 Goldberg D, Bridges K, Duncan-Jones P et al. Detecting anxiety and depression in general medical settings. *BMJ* 1988;297:897-9.
32 Bültmann U, Kant I, Kasl SV et al. Fatigue and psychological distress in the working population: psychometrics, prevalence, and correlates. *J Psychosom Res* 2002;52:445-52.
33 Terluin B. The Four Dimensional Symptom Questionnaire (4DSQ) [In Dutch: De Vierdimensionele Klachtenlijst (4DKL) in de huisartspraktijk]. *De Psycholoog* 1998;33:18-24.
34 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.
35 van Rhenen W, van Dijk FJ, Schaufeli WB et al. Distress or no distress, that's the question: A cutoff point for distress in a working population. *J Occup Med Toxicol* 2008;3:3.
36 Stevens JP. Exploratory and confirmatory factor analysis. Applied multivariate statistics for the social sciences. Mahwah, NJ: Lawrence Erlbaum, 2002: 385-469.
37 Tabachnick BG, Fidell LS. Principal components and factor analysis. 4th ed. Boston: Allyn and Bacon, 2001.
38 Catell RB. The scree test for number of factors. *Multivar Behav Res* 1966;1:245-76.
39 Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas* 1960;20:141-51.
40 Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika* 1965;30:179-85.
41 Ruiz MA, Pardo A, Rejas J et al. Development and validation of the "Treatment Satisfaction with Medicines Questionnaire" (SATMED-Q). *Value Health* 2008;11:913-26.
42 Nunnally JO. Psychometric theory. New York: McGraw Hill, 1978.
43 Stuive I, Kiers HAL, Timmerman ME. The empirical verification of an assignment of items to subtests: the oblique multiple group method versus the confirmatory common factor method. *Educ Psychol Meas* 2008;68:923-39.
44 Stuive I, Kiers HAL, Timmerman ME. Comparison of Methods for adjusting incorrect assignment of items to subtests: Oblique Multiple Group Method. *Educ Psychol Meas* 2009;69:948-65.
45 Watkins MW. Determining parallel anlaysis criteria. *JMASM* 2006;5:344-6.
46 Kaiser HF. A second Generation of Little Jiffy. *Psychometrika* 1970;35:401-15.
47 Kaiser HF. An index of factorial simplicity. *Psychometrika* 1974;39:31-6.
48 Bartlett MS. A note on the multiplying factors for varous chi square approximation. *J R Statist Soc B* 1954;16 (Series B):296-8.

SUPPLEMENT

Supplementing file 1

Results of exploratory factor analysis by Principal Component Analysis (PCA) with varimax rotation per cluster

Table 1. Exploratory factor analysis by PCA with varimax rotation with selected items of cluster 1.

Items	Factors		
	CA	WR	DM
Cognitive aspects of task execution (CA)			
keeping sufficient overview	**0.86**	0.12	0.13
working efficiently	**0.86**	0.11	0.14
carrying out work in general	**0.82**	0.12	0.09
working accurately	**0.81**	0.11	0.21
not forgetting something	**0.81**	0.11	0.12
staying alert	**0.75**	0.21	0.13
working independently	**0.74**	0.04	0.12
Withdrawing from responsibilities (WR)			
avoiding meetings and evaluations	0.03	**0.65**	0.01
preferring routine jobs	0.06	**0.64**	0.04
avoiding supervision of interns	-0.15	**0.58**	-0.12
ducking out of particular tasks	0.05	**0.51**	0.05
avoiding tasks with a lot of responsibility	0.23	**0.51**	0.17
let colleagues make decisions	0.19	**0.50**	0.16
reducing work to what is absolutely necessary	0.18	**0.48**	0.00
avoid decision making	0.13	**0.45**	0.16
Impaired decision making (DM)			
being able to make decisions quickly *(reversed)*	0.12	0.05	**0.91**
being able to make important decisions *(reversed)*	0.13	0.07	**0.89**
being able to make right decisions in stressful situations *(reversed)*	0.18	0.01	**0.81**
being able to make quick decisions in unexpected situations *(reversed)*	0.13	0.13	**0.33**

Note. The highest loading of each item is printed in bold.

Table 2. Exploratory factor analysis by PCA with varimax rotation with selected items of cluster 2.

Items	Factor
	CI
Causing incidents at work (CI)	
working carefully	**0.77**
error-free medication administering/handing	**0.76**
preventing incidents	**0.74**
near incidents in performing and reporting actions	**0.68**
almost causing incidents	**0.62**
making mistakes more often	**0.58**
causing incidents in performing and reporting actions	**0.58**
near mistakes in administration/handing of medicines	**0.56**
making mistakes initiation infusions	**0.49**
underestimate the seriousness of a situation	**0.49**
making mistakes in assessing needed (nursing) care	**0.31**

Note. The highest loading of each item is printed in bold.

Table 3. Exploratory factor analysis by PCA with varimax rotation with selected items of cluster 3.

Items	Factors		
	AC	CA	CP
Avoiding contact with colleagues (AC)			
avoiding common areas	**0.84**	0.16	0.02
avoiding conversations with colleagues	**0.80**	0.26	0.18
leaving for a moment to be alone	**0.78**	0.20	0.11
avoiding working together	**0.75**	0.27	0.04
Conflicts and irritations with colleagues (CA)			
frictions with co-workers	0.10	**0.70**	0.08
tensions between you and colleagues	0.28	**0.69**	0.04
react irritated towards colleagues/ manager	0.16	**0.67**	0.06
often getting angry about matters at work	0.20	**0.66**	0.12
often having conflicts with manager	0.14	**0.59**	0.02
being irritated during work	0.17	**0.58**	0.42
getting upset at work easily	0.31	**0.52**	0.26
Impaired contact with patients and their family (CP)			
showing sufficient empathy	0.07	0.01	**0.84**
contact with patients and family	0.03	-0.08	**0.81**
taking time for patients	0.09	0.07	**0.74**
competently handling aggressive patients or family	0.25	0.10	**0.74**
tone of voice is not too friendly	-0.23	0.41	**0.53**
being curt towards patients or their family	-0.25	0.37	**0.48**
not succeeding in listening well to patients	0.08	0.19	**0.46**
treating patients too roughly	0.25	0.19	**0.37**

Note. The highest loading of each item is printed in bold.

Table 4. Exploratory factor analysis by PCA with varimax rotation with selected items of cluster 4.

Items	Factor
	EM
Lack of energy and motivation (EM)	
not looking forward to working day	**0.84**
having difficulty getting trough a working day	**0.80**
needing an extra day of to get through the working week	**0.77**
starting working day moaning and groaning	**0.75**
working without any enthusiasm	**0.75**
losing control over work	**0.66**

Note. The highest loading of each item is printed in bold.

Supplementing file 2

Results of the Oblique Multiple Group Measures analysis

Table 1. Confirmatory factor analysis by Oblique Multiple Group Measures Analysis.

Items		Loadings on subscales						
		1)	2)	3)	4)	5)	6)	7)
1) Cognitive aspects of task execution and general incidents								
1	keeping sufficient overview	**0.70**	0.36	0.43	0.32	0.40	0.43	0.44
2	not forgetting something	**0.68**	0.34	0.37	0.22	0.30	0.41	0.38
3	working efficiently	**0.70**	0.34	0.39	0.25	0.29	0.41	0.30
4	carrying out work in general	**0.68**	0.28	0.35	0.27	0.39	0.43	0.46
5	working independently	**0.60**	0.24	0.24	0.15	0.26	0.33	0.34
6	staying alert	**0.60**	0.26	0.35	0.29	0.24	0.31	0.28
7	working accurately	**0.69**	0.37	0.39	0.23	0.23	0.38	0.28
16	losing control over work	**0.52**	0.27	0.26	0.35	0.41	0.28	0.38
8	preventing incidents	**0.65**	0.25	0.40	0.26	0.41	0.32	0.36
9	working carefully	**0.72**	0.32	0.40	0.26	0.28	0.38	0.35
15	making mistakes more often	**0.41**	0.09	0.27	0.26	0.22	0.21	0.25
2) Imaired decision making								
48	being able to make important decisions	0.31	**0.66**	0.17	0.11	0.17	0.23	0.13
49	being able to make decisions quickly	0.29	**0.72**	0.22	0.12	0.19	0.22	0.15
50	being able to make right decisions in stressful situations	0.25	**0.56**	0.18	0.02	0.12	0.16	0.14
3) Causing incidents at work								
14	error-free medication administering/handing	**0.67**	0.27	0.29	0.22	0.21	0.29	0.28
26	almost causing incidents	0.31	0.15	**0.42**	0.20	0.25	0.24	0.21
27	near mistakes in administration/handing of medicines	0.28	0.22	**0.35**	0.16	0.21	0.17	0.15
28	underestimate the seriousness of a situation	0.27	0.07	**0.29**	0.12	0.18	0.23	0.16
29	making mistakes initiation infusions	**0.32**	0.21	0.30	0.20	0.19	0.26	0.25
30	making mistakes in assessing needed (nursing) care	0.23	0.11	**0.29**	0.09	0.24	0.19	0.15
31	near mistakes in administration/handing of medicines	0.32	0.29	**0.35**	0.24	0.29	0.27	0.22
32	near incidents in performing and reporting actions	0.40	0.22	**0.43**	0.31	0.26	0.30	0.22
4) Avoidance behavior								
36	leaving for a moment to be alone	0.42	0.07	0.33	**0.46**	0.32	0.22	0.34
37	avoiding conversations with colleagues	0.37	0.04	0.26	**0.58**	0.33	0.17	0.34
38	avoiding common areas	0.31	0.01	0.21	**0.57**	0.30	0.13	0.31
39	avoiding working together	0.22	0.04	0.19	**0.51**	0.33	0.07	0.26
40	let colleagues make decisions	0.23	0.09	0.17	**0.28**	0.20	0.17	0.21
41	reducing work to what is absolutely necessary	0.19	0.14	0.16	0.28	0.18	0.19	**0.29**
42	avoiding meetings and evaluations	0.17	0.11	0.04	0.22	**0.23**	0.11	0.21
43	preferring routine jobs	0.19	0.16	0.19	0.32	0.24	0.24	**0.35**
5) Conflicts and irritations with colleagues								
44	getting upset at work easily	0.30	0.14	0.28	0.33	**0.45**	0.30	0.42
45	react irritated towards colleagues/ manager	0.22	0.12	0.18	0.22	**0.50**	0.18	0.36
46	often getting angry about matters at work	0.30	0.17	0.22	0.25	**0.51**	0.24	0.40
47	often having conflicts with manager	0.36	0.13	0.15	0.26	**0.37**	0.18	0.33
33	frictions with co-workers	0.29	0.16	0.24	0.18	**0.47**	0.24	0.27
34	being irritated during work	0.45	0.28	0.32	0.31	**0.53**	0.35	0.45
35	tensions between you and colleagues	0.27	0.14	0.21	0.33	**0.44**	0.22	0.33
6) Impaired contact with patients and their family								
24	tone of voice is not too friendly	0.24	0.16	0.23	0.07	0.25	**0.43**	0.23
25	being curt towards patients or their family	0.15	0.07	0.22	0.05	0.17	**0.36**	0.17
22	not succeeding in listening well to patients	0.22	0.16	0.28	0.20	0.23	**0.33**	0.26
23	treating patients too roughly	0.17	0.11	0.15	**0.24**	**0.24**	**0.24**	**0.24**
10	contact with patients and family	**0.53**	0.34	0.25	0.13	0.23	0.41	0.28
11	showing sufficient empathy	0.42	0.30	0.29	0.16	0.23	**0.46**	029
12	taking time for patients	**0.62**	0.27	0.28	0.20	0.33	0.44	0.37
13	competently handling aggressive patients or family	**0.48**	0.24	0.25	0.24	0.28	0.40	0.33

Note. The highest loading of each item is printed in bold. Item numbers of final questionnaire are used.

Table 1. (continued)

Items		Loadings on subscales						
		1)	2)	3)	4)	5)	6)	7)
7) Lack of energy and motivation								
20	having difficulty getting trough a working day	0.35	0.13	0.19	0.35	0.37	0.24	**0.57**
21	needing an extra day of to get through the working week	0.41	0.11	0.23	0.33	0.37	0.28	**0.53**
19	working without any enthusiasm	0.28	0.25	0.19	0.33	0.37	0.25	**0.38**
17	not looking forward to working day	0.46	0.22	0.29	0.30	042	0.32	**0.59**
18	starting working day moaning and groaning	0.23	-0.01	0.12	0.13	0.29	0.24	**0.50**

Note. The highest loading of each item is printed in bold. Item numbers of final questionnaire are used.

4

Psychometric properties of the Nurses Work Functioning Questionnaire (NWFQ)

Fania R. Gärtner
Karen Nieuwenhuijsen
Frank J.H. van Dijk
Judith K. Sluiter

PLoS One 2011;6(11):e26565.

ABSTRACT

Objectives: The Nurses Work Functioning Questionnaire (NWFQ) is a 50-item self-report questionnaire specifically developed for nurses and allied health professionals. Its seven subscales measure impairments in the work functioning due to common mental disorders. Aim of this study is to evaluate the psychometric properties of the NWFQ, by assessing reproducibility and construct validity.

Methods: The questionnaire was administered to 314 nurses and allied health professionals with a re-test in 112 subjects. Reproducibility was assessed by the intraclass correlations coefficients (ICC) and the standard error of measurement (SEM). For construct validity, correlations were calculated with a general work functioning scale, the Endicott Work Productivity Scale (EWPS) (convergent validity) and with a physical functioning scale (divergent validity). For discriminative validity, a Mann-Whitney U test was performed testing for significant differences between subjects with mental health complaints and without.

Results: All subscales showed good reliability (ICC: 0.72–0.86), except for one (ICC: 0.16). Convergent validity was good in six subscales, correlations ranged from 0.38-0.62. However, in one subscale the correlation with the EWPS was too low (0.22). Divergent validity was good in all subscales based on correlations ranged from (-0.06) – (-0.23). Discriminative validity was good in all subscales, based on significant differences between subjects with and without mental health complaints ($p < 0.001$ – $p = 0.003$).

Conclusion: The NWFQ demonstrates good psychometric properties, for six of the seven subscales. Subscale “impaired decision making” needs improvement before further use.

INTRODUCTION

Work is one of the most important foundations on which a person's life and quality of life is built, next to health, family and social environment.[1] Not only is work an important source of financial income, but it is also important for identity and self-actualization. Thus, functioning well at work is a necessity for well-being. One known factor that threatens good work functioning is health problems.[2] In particular, mental health problems can negatively impact work functioning and are known to be highly prevalent in the working population.[3-6] In the working population, the prevalence of psychological distress is 23%.[7] In some occupations, the presence of impaired work functioning demands special attention, such as in the health care sector. First, common mental disorders (CMDs) are more prevalent in this sector than in other (service) sectors.[8] Second, impairments in work functioning in this sector can be serious and are not limited to the employee and the organization, as they present severe risks for patients as well.[9 10]

Since the last decade, occupational health psychology and occupational medicine have focused more and more on impairments in work functioning due to health problems, which is also referred to as presenteeism.[2-4 11-18] To gain additional insights into this concept, its causes and its effects, a number of measurement instruments have been developed.[6 19] In these instruments, impaired work functioning (due to health problems) is operationalized differently. Some, like the Work Productivity and Activity Impairment Questionnaire (WPAI), quantify it as hours or days being present at work, but with impaired functioning.[20] Others focus on the work roles, like the multidimensional Work Limitation Questionnaire (WLQ),[21] which differentiates the various aspects of work, i.e., time management demands, physical demands, mental and interpersonal demands, and output demands. What most instruments have in common is that they are designed to be generic, which allows them to be used in various different work settings. One newly developed instrument that distinguishes itself from the existing scales is the Nurses Work Functioning Questionnaire (NWFQ), which we have developed for nurses and allied health professionals.[22] Because of its job-specific nature, it better connects to the work context of nurses and allied health professionals. Its seven subscales capture domains of work that are relevant for these occupations. The job-specificity allows for the items to explicitly describe the concrete experiences and tasks of the work of nurses and allied health professionals. This characteristic of the items facilitates reflection on situations at work and enables self-report. It should be noted that NWFQ scores do not include an overall score of work functioning. Rather, the NWFQ gives insight into various aspects of the work of nurses and allied health professionals that might be impaired due to mental health complaints. The focus on work impairments related to mental health complaints is chosen as mental health complaints are expected to have, at least partly, different effects on work functioning than other (physical) health complaints, e.g., musculoskeletal disorders. For example, we know that mental health complaints can cause cognitive impairments, as the

inability to concentrate can be one of the symptoms of impaired mental health. Therefore, we assume that different types of health complaints ask for different impaired work functioning questionnaires. Unlike other existing instruments measuring health-related work functioning, the items of the NWFQ do not explicitly refer to (known) health problems like the Work Limitation Questionnaire (WLQ) and the Stanford Presenteeism Scale (SPS) do.[21] [23] Three features of the NWFQ contribute to the usefulness on detecting of individuals with work functioning problems due to CMDs and of identifying the specific aspects of impaired work; thus, allowing for purposeful interventions. First, the job-specificity of the NWFQ items, second, the fact that these items do not refer to known health complaints and third the distinction of seven specific aspects of the work make detection of new cases by the NWFQ possible. Interventions that may be initiated based on the NWFQ scores might directly target the work e.g., temporary reorganization of work or discouraging the exertion of specific tasks, but might also address the employee's functioning and mental health complaints through guidance, support, or medical treatment by a (occupational) health professionals.

In an earlier study, some psychometric properties of the NWFQ were already evaluated i.e., the content validity, factorial validity, and the internal consistency.[22] It has been shown that the NWFQ has high content validity, its subscales and items were evaluated as being comprehensive and relevant, and all subscales had acceptable to good internal consistency. Furthermore, its structural validity was good, as the subscale distribution was validated in a confirmatory factor analysis. However, other psychometric properties need to be evaluated further.

Regarding the reproducibility, which is the ability of a measurement tool to reproduce similar results in repeated measures of (stable) subjects, two aspects were evaluated in this study: the level of agreement and the test-retest reliability.[24] The level of agreement gives insight into the stability of the repeated scores within subjects. The test-retest reliability gives an indication of how well subjects can be distinguished from each other despite measurement errors.

In the present study, we also evaluated three types of construct validity.[25] First, we assessed convergent validity, which refers to the relationship between the tested instrument and instruments that measure related constructs. We chose to assess the relationship of the NWFQ with a generic work functioning questionnaire, as we expect them to overlap given that they both assess functioning at work. As the underlying construct of the NWFQ is "impaired work functioning due to mental health complaints", we expect the NWFQ scores to be related to mental health problems. Therefore, we also assessed the relationship between the NWFQ-scores and the mental health complaints for the convergent validity. Second, we evaluated divergent validity to test for the non-relatedness with a measure of a dissimilar construct. For this purpose, we examined the association of the NWFQ subscales with a physical functioning scale, assuming that impairments in work functioning measured by the NWFQ are not related to pure physical health problems. Third, the discriminative validity was studied. As mental

health problems are a probable cause of impairments in the work functioning, we expected to see differences between the groups of workers with and without mental health complaints.

In sum, the aim of this study was to assess the reproducibility of the NWFQ as well as its construct validity, encompassing convergent, divergent, and discriminative validity.

METHODS

Design

This study holds a within subject design with two measurement points: T1 and T2. The data from the first sample at T1 were used for the assessment of the convergent and discriminative validity. The data from the second sample at T2 were used for the divergent validity analysis, and the T2 data combined with the T1 data were used for the reproducibility analyses. The time interval between T1 and T2 was ten to 17 days, as during this span of time workers were expected to be stable with regard to work functioning and mental health.

Subjects

A random sample of 1,200 nurses and allied health professions were contacted in one Dutch academic medical center in order to recruit 300 respondents. The expected response rate was low (25%) because of the large number of items in the questionnaire. The sample was stratified by occupation, gender, and age, and it was representative of the source population, which comprised all employed nurses (including surgical nurses and anesthetic nurses) and allied health professionals of that medical center. The sample at T2 consisted of the 300 employees who first completed the questionnaire at T1. We aimed to recruit 100 respondents for T2.

Procedure

Data collection took place in August and September 2009. Prior to the distribution of the self-administered online questionnaire, the team managers of the relevant departments received information regarding the purpose, aim, and procedure of the study. One week in advance of the distribution, all 1,200 eligible subjects were provided with general information about the study and its purpose through email. Two reminders were sent by email. The first 300 respondents of the questionnaire at T1 were emailed with a request to take part in the retest two weeks after they completed the questionnaire. After one week, a reminder was sent to these 300 subjects. Subjects were provided with an individual username and password to log in at the website with the questionnaire. Agreeing with the informed consent, which was shown online prior to the questionnaire, was a prerequisite for starting the questionnaire. Thus, all participants gave informed consent to participate in the described study.

It was possible to log out halfway through the survey and continue after logging in again. However, the questionnaire had to be fully completed within three days. It was not possible to skip questions. For each filled out questionnaire, we donated 2.50 Euro to a charity that the respondents could select from among three options.

The Medical Ethics Board of the Academic Medical Center Amsterdam gave exemption for ethical approval for the study.

Instruments

Nurses Work Functioning Questionnaire (NWFQ)

The questionnaire tested in this study is the NWFQ developed by Gärtner and colleagues.[22] The NWFQ aims to measure impaired work functioning due to CMDs in nurses and allied health professionals. This 50-item self-report questionnaire consists of seven subscales: 1) *Cognitive aspects of task execution and general incidents*, 2) *Impaired decision making*, 3) *Causing incidents at work* (not suitable for allied health professionals), 4) *Avoidance behavior*, 5) *Conflicts and annoyances with colleagues*, 6) *Impaired contact with patients and their family*, and 7) *Lack of energy and motivation*. Cronbach's alphas vary between 0.70 and 0.94. For the alpha values per scale, see results section. All items of the NWFQ have a reference period of four weeks. Response formats vary between 5-category and 7-category scales; however, the number of categories is the same for all items of one subscale. The content of the response scales varies between Likert-type scales (0 = *totally disagree* to 6 = *totally agree*; 0 = *disagree* to 4 = *agree*; 0 = *no difficulty* to 6 = *great difficulty*), relative frequency categories (0 = *almost never* to 6 = *almost always*; *0 = almost never* to 4 = *almost always*), and absolute frequency categories (0 = *not once* to 6 = *in general more than once a day*). In the present study, in addition to the specific response format for each item, a response category of '*does not apply to my job*' was also provided. In the calculation, this answer was treated as a missing value. The sum scores of the subscales ranged from 0-100 and were calculated as follows: (sum of item scores*100) / (number of items of the subscale*maximum item score). For a more complete description of the development of the questionnaire, the content validity, and the factorial structure, see Gärtner et al.[22]

Endicott Work Productivity Scale (EWPS)

This general work functioning scale is a 25-item self-report questionnaire with a five-point response scale (1 = *never*, 5 = *almost always)*. The sum score is calculated following the traditional scoring method (0, 1, 2, 3, 4), ranging from 0 (best possible score) to 100 (worst possible score). The EWPS is valid and reliable, with a test-retest reliability (10 days to 2 weeks) of ICC = 0.92 and an internal consistency of $\alpha = 0.92$.[26]

General Health Questionnaire (GHQ-12)

This self-report questionnaire was developed to detect common mental disorders in the general population.[27] Its 12 items have a four-point response scale corresponding to the symptoms present (1 = *not at all*, to 4 = *much more than usual*). The reference period for these items was "*the past days*". For the sum score calculation, the traditional GHQ scoring method was used (0, 0, 1, 1), with a range of 0-12. Following earlier studies in working populations, a cut-off point of ≥4 was applied to identify individuals reporting sufficient psychological distress to be classified as probable cases of minor psychiatric disorder.[28]

Four-Dimensional Symptoms Questionnaire (4DSQ)

Of the 4DSQ, which was developed to assess CMDs, the 16-item distress subscale was used.[29 30] The 4DSQ had a reference period of "*the past week*" and a five-point response scale (0 = *no*, 4 = *very often*). The internal consistency of the distress subscale in a working population was $\alpha = 0.90$.[31] For the sum score calculation, we followed the traditional scoring method (0, 1, 2, 2, 2) to generate a continuous distribution ranging from 0-32, where high scores indicated higher stress complaints. For case identification, a cut-off point of ≥11 was applied.[32]

SF-36 physical functioning subscale

This physical functioning self-report scale contained ten items with a three-point response format (1 = *yes, limited a lot*, 2 = *yes, limited a little*, 3 = *no, not limited at all*).[33 34] A sum score was calculated to generate a continuous distribution ranging from 0 (worst health status) to 100 (best health status), using the formula (((sum of raw scores - 10) / 20) * 100). The SF-36 is valid and reliable; for the physical functioning scale, the internal consistency is $\alpha = 0.93$, and the test-retest reliability (2 weeks) is $r = 0.81$.[35]

Demographic data were obtained for each employee. We assessed gender, age, family situation, ethnical background, occupation, number of work hours, labor contract, and years of work experience.

Psychometric analyses

Test-retest reproducibility

For the test-retest reproducibility, we analysed the level of agreement and the test-retest reliability. For the level of agreement we assessed the absolute measurement errors. Therefore, we calculated the standard error of measurement (SEM) for each subscale.[24] The SEM equals the square root of the error variance of an ANOVA analysis, including systematic differences: $SEM = \sqrt{(\sigma^2_{time} + \sigma^2_{error})}$. The SEM values indicate that if a within-subject comparison is made for two sum scores on a subscale of the NWFQ at different points in time, and a change score is smaller than the SEM, then it should be considered a measurement

error. To visualize the level of agreement, a Bland and Altman plot with 95% confidence interval was designed.[36 37] These plots show the difference scores of the subject in relation to the mean score of the test and retest. In the plots the mean change score and the 95% limits or agreement are illustrated. To detect possible systematic errors, the Pearson's correlation coefficients of the difference scores and the mean scores are given, as well a t-test is performed, to see if the mean change score significantly differs from zero.[38]

The test-retest reliability evaluates the ability of the NWFQ to distinguish between subjects despite measurement error. Therefore, the intraclass correlations coefficient (ICC) using the T1 and T2 data (N = 212) was computed for all subscales. To determine the ICC a two-way random effects model was used, the ICC(A.1) according to MCGraw and Wong.[39] The ICC calculation method in which systematic differences are considered to be part of the measurement error was used, called the ICC absolute agreement. The formula used was: ICC = σ^2_p / (σ^2_p + σ^2_{time} + σ^2_{error}).[24] For the ICC, we expected a minimum of 0.70 as sufficient for good reliability.[37]

An assumption in reproducibility analyses is that the sample used is stable regarding the studied concept.[37] We expected our sample to be stable during the two weeks interval. To control for stability, we asked the participants: "*Did your state of well-being change after you first filled out our questionnaire?*" at the second measurement point. Subjects who answered "*no*" were regarded as stable subjects. The level of agreement and reliability analyses were performed separately for the whole sample and the sample with the stable subjects only. However, conclusions were based on the results of the stable sample only.

Construct validity

Three types of construct validity were assessed, convergent validity, divergent validity, and discriminative validity. The convergent validity was assessed by calculating correlations between the NWFQ subscales and the EWPS. As the NWFQ data were not normally distributed, Spearman correlations were used. For good convergent validity, we expected moderate (> 0.30 ≤ 0.60) to high (> 0.60) correlations in a positive direction for the relationship with the EWPS.[40]

The divergent validity was assessed by calculating Spearman correlations between the NWFQ and the SF36 physical functioning scale. We chose the physical functioning subscale because its construct is dissimilar to the construct of the NWFQ, though it is not completely unrelated as both constructs refer to functioning of subjects. For good divergent validity, we expected these correlations to be low (≤ 0.30).[40]

Discriminative validity was assessed to evaluate the ability of the NWFQ subscales to discriminate between groups that were expected to differ. Therefore, we used a Mann-Whitney U test to test for significant differences in NWFQ scores between workers with and without mental health complaints. Having mental health complaints was defined as scoring above the cut-off on one or both of the mental health complaints scales (GHQ-12 and 4DSQ-distress). To

correct for the high number of tests performed (one for each of the seven subscales), we used a Bonferroni adjustment. Therefore, $p < 0.007$ was regarded as significant.

RESULTS

Of the 1,200 nurses and allied health professionals invited, 314 employees fully completed the questionnaire at the first measurement point (26% response rate). Of these 314, 112 (36%) completely filled out the second questionnaire (T2). Table 1 presents the socio-demographic characteristics of the samples at T1 and T2.

Reproducibility

Twenty-eight subjects responded that their well-being improved or deteriorated between the two measurement points. Therefore, we did not regard them as stable enough to include in our reproducibility analyses.

Level of agreement

The SEM ranged from 2.95 to 6.12 for the six subscales, and there was one outlier with a SEM of 17.11 for the subscale 2) *Impaired decision making* (Table 2). Additionally, the Bland and Altman plots are shown for each NWFQ subscale, based on the stable sample (Figure 1 –

Table 1. Socio-demographic characteristics of the two samples.

Demographic characteristics	T1 (N = 314)		T2 (N = 112)	
Gender (N (%))				
Female	257	(81.8)	94	(83.9)
Male	57	(18.2)	18	(16.1)
Age in years (mean (SD))	44.5	(12.0)	46.3	(10.5)
Marital status (N (%))				
Married/ living together with a partner	227	(72.3)	84	(75)
In a relationship	21	(6.7)	8	(7.1)
Single	54	(17.2)	15	(13.4)
Divorced	11	(3.5)	5	(4.5)
Widow/ widower	1	(0.3)	0	(0)
Ethnical background (N (%))				
Dutch	261	(83.1)	94	(83.9)
Immigrant first generation	35	(11.1)	13	(11.6)
Immigrant second generation	18	(5.7)	5	(4.5)
Occupation (N (%))				
Nurse	220	(70.1)	74	(66.1)
Surgical nurse	23	(7.3)	7	(6.2)
Anesthetic nurse	13	(4.1)	6	(5.4)
Allied health professional	58	(18.5)	25	(22.3)
Working experience in years (mean (SD))	20.8	(12.2)	22.3	(11.6)
Labor contract (N (%))				
Permanent position	301	(95.9)	107	(95.5)
Fixed-term contract	9	(2.9)	1	(0.9)
Temporary employment	4	(1.3)	4	(3.6)
Work hours per week (mean (SD))	30	(6.3)	29	(6.8)

Figure 7). The Bland and Altman plots show the difference in NWFQ scores at the two measurement points. The dots in the figure present the difference scores of the subjects and the lines picture the 95% confidence interval. The 95% confidence intervals vary between the subscales, for four subscales they range from about -8 to 13, for two subscales they are a little larger about -14 to 18. For subscale 2) *Impaired decision making*, the 95 % confidence interval is very large, with -44 to 51. In all subscales except for subscale 5) *Conflicts and annoyances with colleagues*, the mean change score is close to zero and no significant correlation between the mean scores of T1 and T2 and the difference scores are found. In subscale 5) *Conflicts and annoyances with colleagues*, the mean difference score of 2.18 statistically differs from zero with a p-value of 0.018, the correlation coefficient of the mean scores of T1 and T2 and the difference scores is 0.379.

In the plot of subscale 2) *Impaired decision making*, several high change scores can be seen, one is up to 100% of the scale with a change score of 100. Also for this subscale the confidence interval is high, its values are nearly half of the scale range with 51.16 and -44.08.

Test-retest reliability

Based on the sample with stable subjects, the single measure ICC between measurements at T1 and T2 were good for six of the seven subscales, with a range of 0.72 to 0.88 (Table 2). Subscale 2) *Impaired decision making* had a poor reliability score with an ICC of 0.16. Figure 1 to Figure 7 present the Bland and Altman plots for each subscale.

Construct validity

Convergent validity

The Spearman's rank correlation coefficients between the NWFQ subscales and the EWPS sum scale ranged from 0.22 to 0.62 (Table 3). There was one low correlation for subscale 2 ($r = 0.22$), five medium correlations for subscales 3, 4, 5, 6, and 7 and one high correlation for subscale 1 ($r = 0.62$).

Divergent validity

The Spearman's rank correlation coefficients between the subscales of the NWFQ and the SF-36 physical functioning sum score were all low, ranging from -0.23 to -0.06 (Table 3).

Discriminative Validity

Significant differences in the expected direction were found between the group with and without mental health complaints for all seven subscales (Table 4). The p-values ranged from $p < 0.001$ to $p = 0.003$.

Table 2. ICC and SEM of the sample who completed both questionnaires at T1 and T2.

NWFQ subscales	First measure (T1)			Second measure (T2)			Mean difference of T1-T2 (SD)		SEM	ICC	ICC 95% CI	
	Valid N	Median (range)		Valid N	Median (range)						Lower bound	Upper bound
Stable sample (N = 84)												
1. Cognitive aspects of task execution and general incidents	81	3	(0-55)	83	2	(0-73)	1.14	(5.98)	4.27	**0.85**	0.77	0.90
2. Impaired decision making	81	0	(0-100)	83	0	(0-100)	3.54	(24.30)	17.11	0.16	-0.05	0.37
3. Causing incidents at work	41	2	(0-40)	41	0	(0-44)	0.16	(4.36)	3.04	**0.88**	0.79	0.94
4. Avoidance behavior	75	0	(0-53)	74	3	(0-44)	0.14	(4.78)	3.33	**0.79**	0.69	0.87
5. Conflicts and annoyances with colleagues	83	4	(0-50)	83	9	(0-50)	2.18	(8.19)	6.03	**0.72**	0.59	0.81
6. Impaired contact with patients and family	61	2	(0-42)	61	2	(0-48)	0.60	(4.21)	2.95	**0.86**	0.76	0.91
7. Lack of energy and motivation	82	7	(0-63)	84	3	(0-60)	1.18	(8.63)	6.12	**0.74**	0.63	0.83
Total sample (N = 112)												
1. Cognitive aspects of task execution and general incidents	107	5	(0-55)	111	3	(0-73)	0.20	(8.78)	6.17	**0.70**	0.59	0.79
2. Impaired decision making	109	0	(0-100)	111	0	(0-100)	3.86	(21.90)	15.58	0.32	0.14	0.48
3. Causing incidents at work	58	2	(0-40)	58	2	(0-44)	0.15	(5.04)	3.52	**0.82**	0.71	0.89
4. Avoidance behavior	100	0	(0-53)	100	0	(0-44)	0.27	(6.33)	4.41	0.66	0.52	0.76
5. Conflicts and annoyances with colleagues	111	4	(0-50)	111	4	(0-61)	1.14	(9.50)	6.79	0.67	0.56	0.76
6. Impaired contact with patients and family	80	2	(0-42)	82	2	(0-48)	1.42	(7.54)	5.30	0.60	0.42	0.73
7. Lack of energy and motivation	109	7	(0-63)	111	3	(0-60)	0.76	(8.38)	5.92	**0.78**	0.70	0.85

Bold printed value: value supports the hypotheses.

Table 3. Overview of subscale characteristics and correlations for the construct validity analyses.

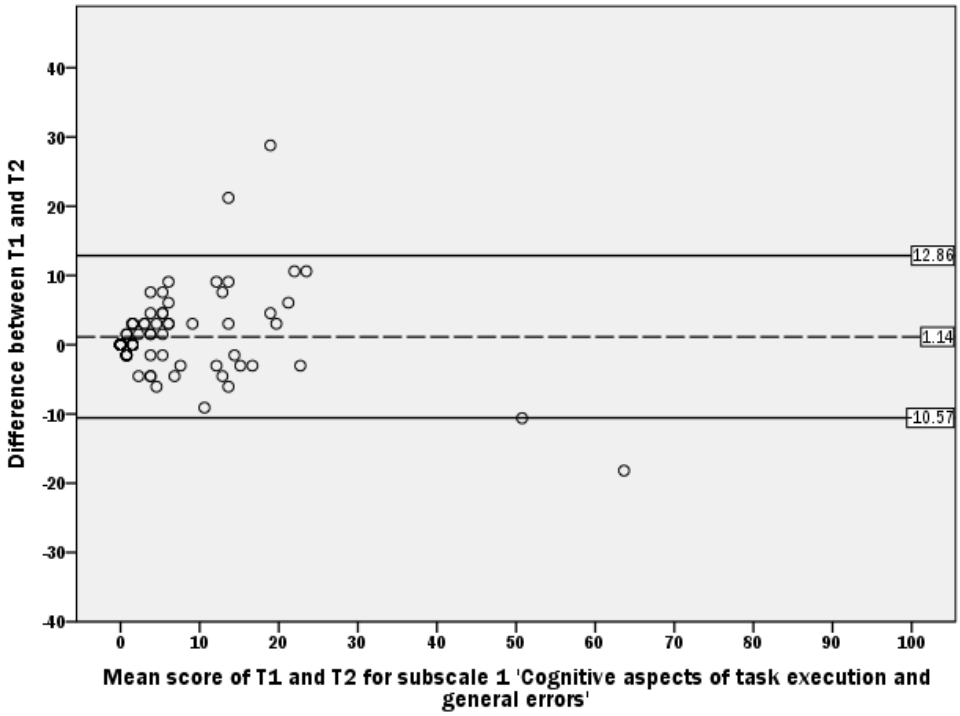
NWFQ subscales	# of items	T1 (total N = 314)				T2 (total N = 112)		
		N	Cronbach's α	Median (range)	Spearman's correlation EWPS	N	Cronbach's α	Spearman's correlation SF-36 physical functioning
1. Cognitive aspects of task execution and general incidents	11	308	0.94	5 (0-82)	**0.62**	113	0.94	**-0.19**
2. Impaired decision making	3	310	0.88	0 (0-100)	0.22	113	0.80	**-0.11**
3. Causing incidents at work	8	178	0.78	4 (0-40)	**0.42**	60	0.88	**-0.06**
4. Avoidance behavior	8	294	0.70	0 (0-81)	**0.38**	102	0.61	**-0.23**
5. Conflicts and annoyances with colleagues	7	311	0.77	4 (0-61)	**0.49**	113	0.74	**-0.11**
6. Impaired contact with patients and family	8	223	0.81	4 (0-42)	**0.50**	83	0.81	**-0.10**
7. Lack of energy and motivation	5	307	0.81	7 (0-73)	**0.53**	113	0.81	**-0.13**

Bold printed value: value supports the hypotheses.

Table 4. Discriminative validity using a Mann-Whitney U test with the T1 sample.

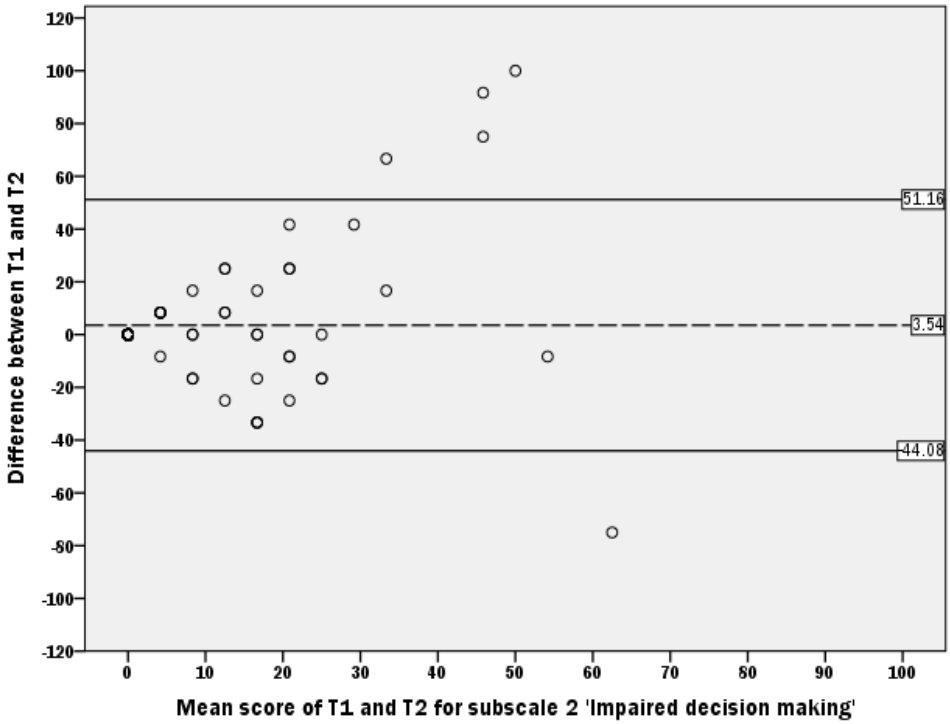
NWFQ subscales	Total			Healthy			CMD*			P-value of Mann-Whitney U test
	N	Median	(range)	N	Median	(range)	N	Median	(range)	CMD*
1. Cognitive aspects of task execution and general incidents	308	5	(0-82)	227	3	(0-39)	81	11	(0-82)	**<0.001**
2. Impaired decision making	310	0	(0-100)	227	0	(0-100)	83	8	(0-92)	**0.003**
3. Causing incidents at work	178	4	(0-40)	130	2	(0-19)	48	6	(0-40)	**0.003**
4. Avoidance behavior	294	0	(0-81)	214	0	(0-31)	80	6	(0-81)	**<0.001**
5. Conflicts and annoyances with colleagues	311	4	(0-61)	228	0	(0-50)	83	14	(0-61)	**<0.001**
6. Impaired contact with patients and family	223	4	(0-42)	167	2	(0-29)	56	6	(0-42)	**0.001**
7. Lack of energy and motivation	307	7	(0-73)	226	3	(0-63)	81	20	(0-73)	**<0.001**

Bold printed value: significant difference between the two groups, above versus below the cut-off score using a Bonferroni-adjusted p-value of p <0.007.
* above cut-off for GHQ-12 and/or 4DSQ-distress



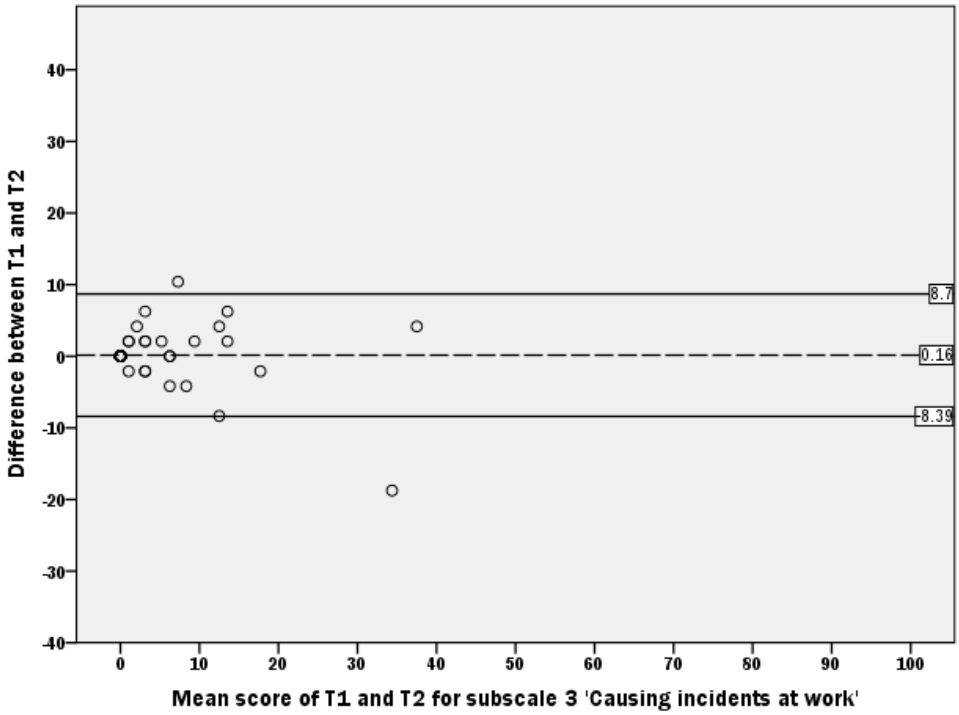
Pearson's correlations coefficient = 0.144
T-test for the difference of the mean change score with 0: p = 0.090

Figure 1. Bland and Altman plot for subscale 1) *Cognitive aspects of task execution and general incidents*.



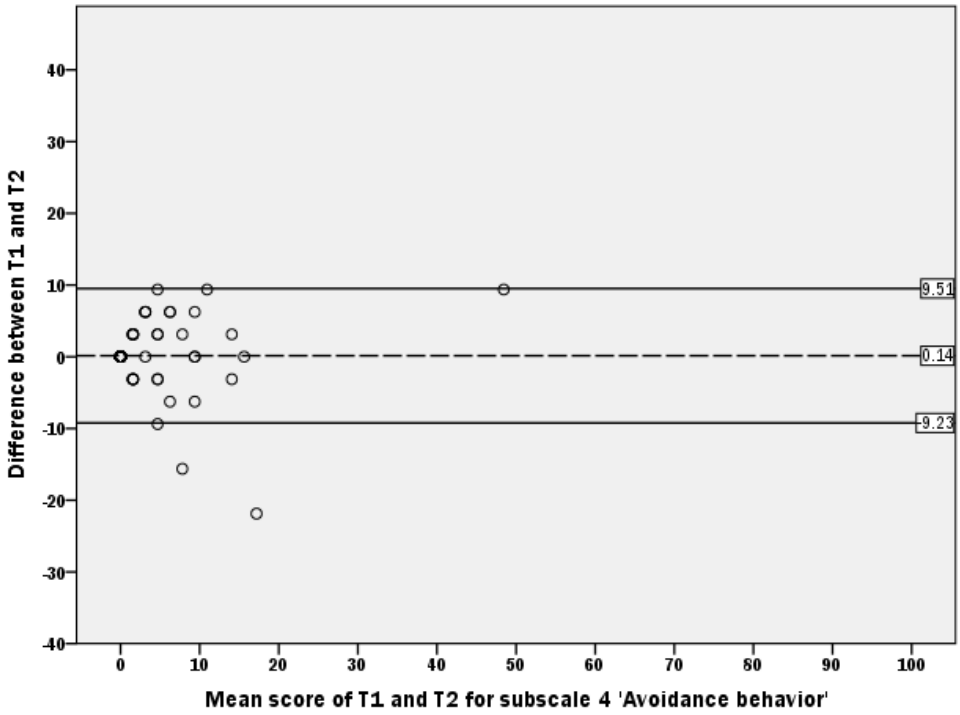
Pearson's correlations coefficient = 0.103
T-test for the difference of the mean change score with 0: p = 0.196

Figure 2. Bland and Altman plot for subscale 2) *Impaired decision making*.



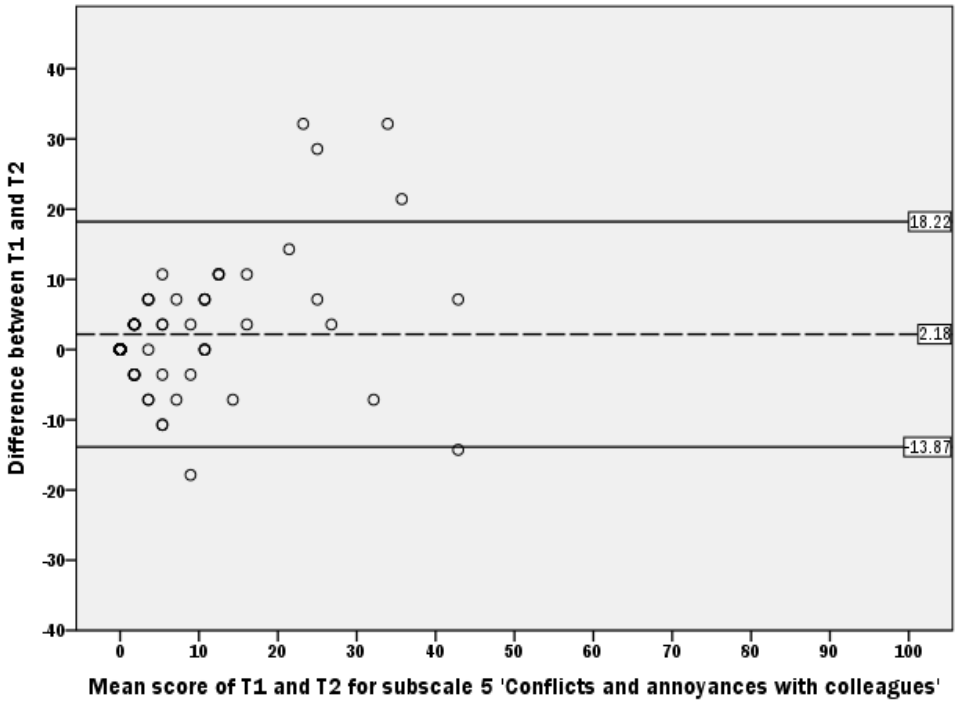
Pearson's correlations coefficient = 0.146
T-test for the difference of the mean change score with 0: p = 0.822

Figure 3. Bland and Altman plot for subscale 3) *Causing incidents at work*.



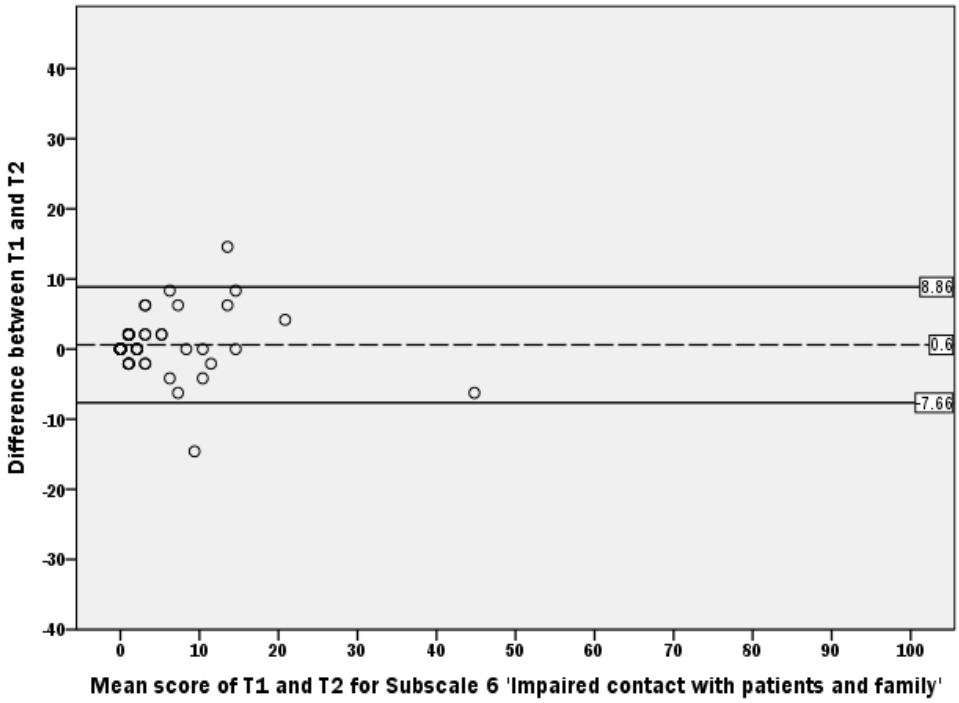
Pearson's correlations coefficient = 0.134
T-test for the difference of the mean change score with 0: p = 0.813

Figure 4. Bland and Altman plot for subscale 4) *Avoidance behavior*.



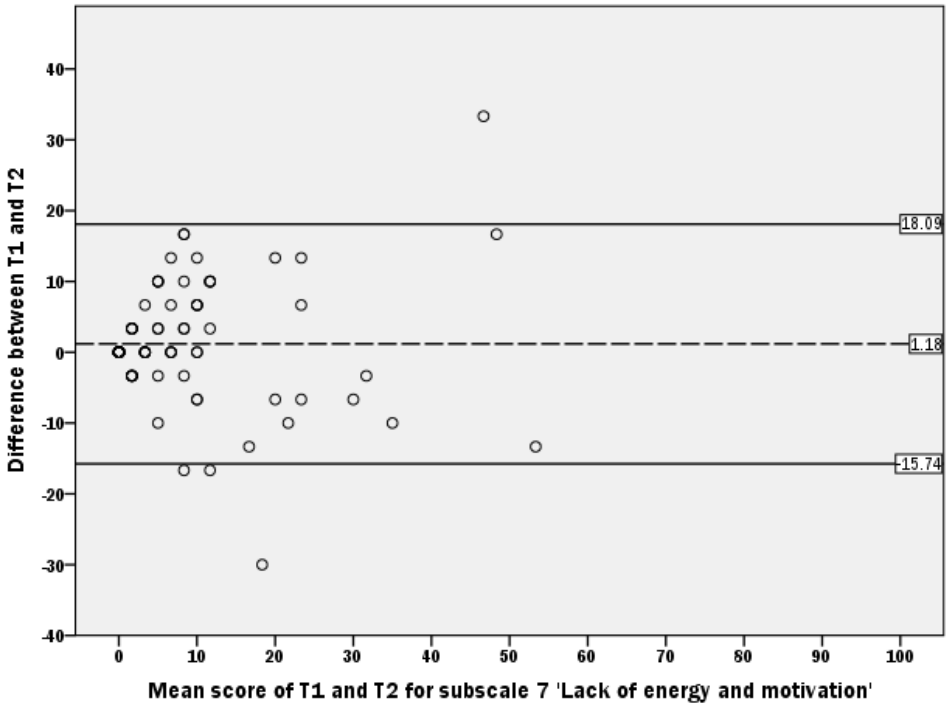
Pearson's correlations coefficient = 0.379
T-test for the difference of the mean change score with 0: p = 0.018

Figure 5. Bland and Altman plot for subscale 5) *Conflicts and irritations with colleagues*.



Pearson's correlations coefficient = 0.097
T-test for the difference of the mean change score with 0: p = 0.309

Figure 6. Bland and Altman plot for subscale 6) *Impaired contact with patients and their family*.



Pearson's correlations coefficient = 0.068
T-test for the difference of the mean change score with 0: p = 0.220

Figure 7. Bland and Altman plot for subscale 7) *Lack of energy and motivation*.

DISCUSSION

The purpose of this study was to evaluate the psychometric quality of the newly developed NWFQ in terms of reliability and construct validity. Overall, the results were satisfactory for six of the seven subscales.

Except for subscale 2) *Impaired decision making*, the subscales of the NWFQ had good reproducibility and thus were able to distinguish between subjects, even when measurement error was taken into account. The SEM values, expressed in the same value as the target instrument, help to interpret the changes in scores of individuals over time on the NWFQ subscales. When within subjects comparisons were made, changes had to be larger than the SEM to ensure that the observed differences were not due to measurement error. Based on the Bland and Altman plots, we can state that level of agreement is good for six of the seven subscales. For subscale 5) *Conflicts and annoyances with colleagues* systematic error appears to influences the score. Based on the dots in the plot, we suppose a possible reason for systematic error might be that subjects with high mean scores at T1 tend to improve at T2, rather than being stable or deteriorate. In future studies on the characteristics of the NWFQ, such as assessment of the responsiveness, these possible systematic differences should be taken into account, e.g., by subgroup analyses in which analyses are stratified for groups of subjects with high and low baseline scores.

Our data offer strong support for good construct validity as the hypothesized relationships were confirmed, with the exception of subscale 2) *Impaired decision making*. Regarding the convergent validity of these six subscales, all correlations with the EWPS were substantial and in line with the hypothesis. The fact that the correlations were medium and not high verifies that the NWFQ, on the one hand, has enough overlap with a generic work productivity scale. On the other hand, this job-specific instrument measures aspects of additional value compared to a generic questionnaire. Regarding divergent validity, the hypothesis that the correlations between the NWFQ and the unrelated physical functioning measure are low is supported for all the subscales. All scales showed clear discriminative validity; thus, they discriminate well between a group of subjects with and without mental health complaints. Therefore, the relatedness of CMDs with impaired work functioning is evident.

It is obvious that subscale 2) *Impaired decision making* performed the weakest in our evaluation of the psychometric properties of the NWFQ. The subscale failed to show good reliability. In a subscale with only three items, small differences in scores on one item have bigger impact for the stability of the measures than in scales with more items. High reliability scores are therefore more difficult to derive in subscales with smaller number of items. However, increasing the number of items is no attractive alternative, as in the development process of the NWFQ, the three item option for this subscale led to the best internal consistency and interpretability. We therefore must conclude that in our sample, the subscale

was not able to distinguish between individuals. In addition, the subscale *impaired decision making*, failed to support the hypothesis for good convergent validity. Consequently, we have to discourage the use of subscale 2) *Impaired decision making* in the present form. However, we still regard impaired decision making to be an important aspect of the construct of the NWFQ. During the development process of the NWFQ, the aspect impaired decision making as an effect of CMDs was discussed repeatedly in focus groups with nurses and professionals. Additionally, in expert checks, the content validity of the subscale was confirmed and impaired decision making was evaluated as an important aspect of the overall construct.

Methodological notes

A methodological limitation of this study that deserves consideration is that the data were primarily collected within the scope of the questionnaire development. The items of the tested subscales were included in the original item pool, which was much longer than the final version of the NWFQ. Therefore, the context of questionnaire administration was not exactly the same as it will be in future use. This is evident in the low response rate for our study, which can partly be explained by the length of the questionnaire administered. However, the sample was representative for the gender and age distribution of the studied occupations in the medical center.

Recommendations for further research

We want to point out three aspects for further research. The first point concerns the subscale 2) *Impaired decision making*. As described above, we regard that subscale as necessary part of the NWFQ; however, no reproducible and valid form of measuring that subconstruct is found yet. Therefore, future research should identify new items measuring impaired decision making in nurses and allied health professionals with CMDs that form a subscale with good psychometric quality. Second, the conclusion drawn from the presented data is only valid for the Dutch version of the NWFQ and for nurses and allied health professionals working in academic medical centers. Although a backward-forward translation of the questionnaire into English exists, we recommend additional evaluations of the psychometric quality of that version. Third, future use of the NWFQ as a diagnostic instrument in occupational health practice, suitable cut-off values for the subscales need to be identified. In addition, the responsiveness, the smallest detectable change (SDC) and minimal important change (MIC), would be important to assess as they would allow for making inferences based on the changes in scores of individual workers on the NWFQ over time.

Recommendations for practice

Until now, work functioning instruments have mainly been used in the scientific setting for research aims. It would be of great value to apply them for use in occupational health practice as well, and in particular, applying the NWFQ for preventive aims would be of value. For

preventive purposes, work functioning instruments must help to realize two aims; first, timely interventions and second, purposeful interventions on mental health complaints and related impairments in the functioning. We conclude that the nature of the NWFQ fulfills both these needs. The NWFQ can be used for detection purposes as its items do not refer to known health problems; furthermore, its multidimensionality makes identification of specific work aspects possible, and therefore is a starting point for purposeful interventions. One possible way to initiate the detection and monitoring of employees with mental health complaints and related work functioning problems for preventive purposes would be in a Workers' Health Surveillance in which the NWFQ could be included.

The NWFQ is available for use, see supporting file S1. Though, users have to follow Creative Commons Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0).

Conclusion

The NWFQ demonstrated good psychometric properties for six subscales. Subscale 2) *Impaired decision making*, did not show enough ability to discriminate between subject and the association with other work functioning measure was too weak; therefore, we discourage use of that subscale in the present form. In conclusion, the NWFQ is a reproducible and valid instrument suitable for the measurement of impairments in work functioning due to CMDs in nurses and allied health professionals when including six of the seven subscales.

REFERENCES

1 Hultman B, Hemlin S, Hornquist JO. Quality of life among unemployed and employed people in northern Sweden. Are there any differences? *Work* 2006;26:47-56.
2 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
3 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
4 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
5 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
6 McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev* 2009;29:243-59.
7 Bültmann U, Kant I, Kasl SV et al. Fatigue and psychological distress in the working population: psychometrics, prevalence, and correlates. *J Psychosom Res* 2002;52:445-52.
8 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.
9 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
10 Suzuki K, Ohida T, Kaneita Y et al. Mental health status, shift work, and occupational accidents among hospital nurses in Japan. *J Occup Health* 2004;46:448-54.
11 Adler DA, McLaughlin TJ, Rogers WH et al. Job performance deficits due to depression. *Am J Psychiatry* 2006;163:1569-76.
12 Aronsson G, Gustafsson K. Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research. *J Occup Environ Med* 2005;47:958-66.
13 Burton WN, Conti DJ, Chen CY et al. The role of health risk factors and disease on worker productivity. *J Occup Environ Med* 1999;41:863-77.
14 Burton WN, Pransky G, Conti DJ et al. The association of medical conditions and presenteeism. *J Occup Environ Med* 2004;46:S38-S45.
15 Haslam C, Atkinson S, Brown S et al. Perceptions of the impact of depression and anxiety and the medication for these conditions on safety in the workplace. *Occup Environ Med* 2005;62:538-45.
16 Pilette PC. Presenteeism in nursing: a clear and present danger to productivity. *J Nurs Adm* 2005;35:300-3.
17 Schultz AB, Edington DW. Employee health and presenteeism: a systematic review. *J Occup Rehabil* 2007;17:547-79.
18 Stewart WF, Ricci JA, Chee E et al. Cost of lost productive work time among US workers with depression. *JAMA* 2003;289:3135-44.
19 Nieuwenhuijsen K, Franche RL, van Dijk FJ. Work functioning measurement: tools for occupational mental health research. *J Occup Environ Med* 2010;52:778-90.
20 Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993;4:353-65.
21 Lerner D, Amick Iii BC, Rogers WH et al. The Work Limitations Questionnaire. *Med Care* 2001;39:72-85.
22 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. Impaired work functioning due to common mental disorders in nurses and allied health professionals: the Nurses Work Functioning Questionnaire. *Int Arch Occup Environ Health* 2011;85:125-38.
23 Koopman C, Pelletier KR, Murray JF et al. Stanford presenteeism scale: health status and employee productivity. *J Occup Environ Med* 2002;44:14-20.
24 de Vet HC, Terwee CB, Knol DL et al. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033-9.
25 Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. fourth ed. Oxford: Oxford University Press, 2008.
26 Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull* 1997;33:13-6.
27 Goldberg D, Bridges K, Duncan-Jones P et al. Detecting anxiety and depression in general medical settings. *BMJ* 1988;297:897-9.
28 Bültmann U, Kant IJ, Kasl SV et al. Lifestyle factors as risk factors for fatigue and psychological distress in the working population: prospective results from the Maastricht Cohort Study. *J Occup Environ Med* 2002;44:116-24.
29 Terluin B. The Four Dimensional Symptom Questionnaire (4DSQ) [In Dutch: De Vierdimensionele Klachtenlijst (4DKL) in de huisartspraktijk]. *De Psycholoog* 1998;33:18-24.

30 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.
31 Terluin B, van Rhenen W, Schaufeli WB et al. The four-dimensional symptom questionnaire (4DSQ): measuring distress and other mental health problems in a working population. *Work Stress* 2004;18:187-207.
32 van Rhenen W., van Dijk FJ, Schaufeli WB et al. Distress or no distress, that's the question: A cutoff point for distress in a working population. *J Occup Med Toxicol* 2008;3:3.
33 Jenkinson C, Layte R, Coulter A et al. Evidence for the sensitivity of the SF-36 health status measure to inequalities in health: results from the Oxford healthy lifestyles survey. *J Epidemiol Community Health* 1996;50:377-80.
34 Ware JE, Jr., Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998;51:903-12.
35 Brazier JE, Harper R, Jones NM et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4.
36 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
37 Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
38 Mantha S, Roizen MF, Fleisher LA et al. Comparing methods of clinical measurement: reporting standards for bland and altman analysis. *Anesth Analg* 2000;90:593-602.
39 Mcgraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996;1:30-46.
40 Innes E, Straker L. Validity of work-related assessments. *Work* 1999;13:125-52.

5

Interpretability of change of the Nurses Work Functioning Questionnaire: minimal important change and smallest detectable change

Fania R. Gärtner
Karen Nieuwenhuijsen
Frank J.H. van Dijk
Judith K. Sluiter

Accepted for publication in the Journal of Clinical Epidemiology.

ABSTRACT

Objectives: The current study evaluates the interpretability of individual changes and assesses the Minimal Important Change for improvement (MIC) and the Smallest Detectable Change (SDC) of the Nurses Work Functioning Questionnaire (NWFQ).

Study design and setting: Data of two time points, baseline and three months follow-up, of a randomized-controlled trial were used. A 358 nurses and allied health professionals of one Dutch academic medical center participated at both measurement points. Separate MIC values were calculated for the total score of the NWFQ and its six subscales, using two anchor-based methods: mean-change method and ROC-curve method. Two methods for baseline corrections were applied: subgroup analyses and MIC calculation based on relative change scores. The SDCs were calculated using the standard error of measurement.

Results: MIC values ranged from 3.4-8.3 for the mean-change method and from 1.5-9.5 for the ROC-curve method. In a subgroup with high baseline scores, MIC values of the two methods ranged from 4.4-29 and 9.5-41.5 respectively. The SDC values ranged from 7.2-17. Only one MIC value exceeded the SDC; however, 10 of the 14 MIC values exceeded the SDC in the high baseline group.

Conclusion: Three out of seven NWFQ scales exhibited sufficient interpretability of individual change. For four scales conclusions on the interpretability of change can not yet be drawn. SDCs were small compared to the scale-range.

INTRODUCTION

Mental health illness impairs the work functioning of workers who maintain working during complaints.[1-5] These impairments in work functioning due to mental health complaints may have serious consequences in some occupations, such as the health care service, with risks for patient safety. Effective interventions that address the mental health complaints and its effect on work functioning would benefit employers and workers. For this aim a few interventions have been developed and evaluated over the last decade. One example is an intervention study of an adjuvant occupational therapy in workers with depression.[6]

The efficacy of interventions in the enhancement of work functioning has primarily been evaluated using self-reported outcome measures. However, systematic reviews reveal that many of the existing instruments of work functioning have shortcomings in quality evaluation, especially in the interpretability of individual change, defined as the ability to detect change in the measured construct over time.[7 8] The Nurses Work Functioning Questionnaire (NWFQ) is a newly developed measurement scale for the assessment of work function impairments due to mental health complaints.[9] The NWFQ is a 47-item self-report questionnaire with six subscales that measure various aspects of work function impairments and that was specifically developed for nurses and allied health professionals. A recent study revealed that the NWFQ is a valid and reliable measurement.[10] However for application in intervention studies and in practice, the interpretability of change requires further investigation in order to facilitate the interpretation of individuals change scores.[11] The statistical significance of differences alone does not provide insight on whether changes in outcome scores are 'real' changes or changes that occur due to measurement error. Moreover, the clinical relevance of 'real' changes remains questionable. The interpretability of change includes two measures, the 'Minimal Important Change' (MIC) and the 'Smallest Detectable Change' (SDC). The MIC value refers to the smallest difference between two scores on a measurement scale that can be regarded as relevant or important.[12 13] The SDC value refers to the smallest change that can be regarded as a real change that is not due to measurement error.[11 14] With these two measures we have an indication on whether the SDC is small enough to detect minimally important changes, MICs.

Both values aid in the interpretation of results of individual subjects for the scientific evaluation of interventions. Furthermore, occupational health practitioners who advice nurses with mental health complaints would receive indications on relevant changes in the work functioning of their patients. Therefore, this study will assess the MIC and SDC of NWFQ total scores and each of the different subscales.

The methodology for MIC calculation has been in development since its introduction in the late 1980's. Jeaschke and colleagues first introduced a definition and method for MIC calculation.[15] Various review articles on potential methods have been published previously,[11-13 16-21] but no consensus on the best methodology has been reached. MIC calculation is

possible for changes in improvement and deterioration. This study will focus on the MIC for improvement, which from this point will be referred to as MIC. Two main approaches, have been described in the literature, the distribution-based methods[22-24] and the anchor-based methods.[15 25] As Terwee and colleagues 2010 define: distribution-based methods use statistical measures as a value for the minimal important change and anchor-based methods use an external criterion, often called "anchor" to define an important change.[21] However, based on a comparison of these two methods, Terwee and colleagues conclude that from a conceptual point of view the anchor-based methods should be preferred above the distribution-based methods.[21] We agree with the preference for anchor-based methods, because on our view, anchors represent perceptions of subjects on relevant changes, while purely statistical methods are barely possible to provide answers on relevance. Therefore, we will apply the anchor-based methods in this study, in which an external standard is used for the MIC calculation using a self-report measure on the perception of an important change as an anchor, as is recommended by Crosby and colleagues.[17] As several MIC calculation methods are available in the anchor-based approach, we applied the triangulation of two methods, the mean-change method and the Receiver Operating Characteristics (ROC)-curve method.

Another issue is the baseline value of subjects. Two methods are mentioned in the literature to take into account differences in the baseline values: the analysis of subgroup MIC values and the calculation of the MIC based on relative change scores.[17 18 21 26] Both methods will be applied in the current study, because no consensus on a preferred method has been reached.

The first objective of this study is the evaluation of good interpretability of individual change scores by calculating the Area Under the Curve (AUC) of the ROC curve and by comparing the resulting MIC values with the SDC value. Second objective is to calculate the MIC and the SDC for the interpretation of individual change scores in research and practice. These values will be calculated for the NWFQ total score and for the separate NWFQ subscales.

METHODS

Design

This study used a repeated measures design with a baseline (T0) and a three-month follow-up measure (T1). Data collection was integrated in the Mental Vitality @ Work Study, a cluster randomized controlled trial (RCT) with three arms: one control group and two intervention groups. The RCT examines the effect of a workers' health surveillance mental module for nurses and allied health professionals. Participants in both intervention arms completed an online screening questionnaire to detect problems in mental health and work functioning, and

received immediate feedback on their screening results. A consultation with an occupational physician was offered in cases of impairments in mental health or work functioning in the first intervention arm. The second intervention arm offered self-help e-mental health interventions. The study design and procedure have been described previously by Gärtner and colleagues.[27] The data obtained in all three study arms were used in the present study. The Medical Ethics Committee of the Academic Medical Center Amsterdam approved the study protocol.

Subjects and procedure

Participant recruitment occurred in March 2011 in one Dutch academic medical center. All 557 nurses and allied health professionals that participated in the trial were eligible for the present study. Baseline and follow-up questionnaires were provided online. All potential participants received an invitation in their work e-mail account. Participation in the study was voluntary, and informed consent was required prior to the initiation of the baseline questionnaire.

Measurements

The following demographic and occupational characteristics were assessed for descriptive purposes: gender, age, marital status, ethnicity, occupation, years of experience, working hours per week, and type of contract.

Nurses Work Functioning Questionnaire

The Nurses Work Functioning Questionnaire (NWFQ) was administered at baseline and three-month follow-up. The NWFQ is a self-report questionnaire that measures impaired work functioning due to common mental disorders in nurses and allied health professionals.[9 10] Six NWFQ subscales with a total of 47 items were assessed: *Cognitive aspects of task execution and general incidents* (11 items); *Causing incidents at work* (8 items) (not applicable for allied health professionals); *Avoidance behavior* (8 items); *Conflicts and annoyances with colleagues* (7 items); *Impaired contact with patients and their family* (8 items); and *Lack of energy and motivation* (5 items). Cronbach's alphas vary between 0.70 and 0.94. Response formats vary between 5- and 7-category scales. The content of the response scales varies between Likert-type scales (from 0 = *totally disagree* to 6 = *totally agree*), relative frequency categories (from 0 = *almost never* to 6 = *almost always*), and absolute frequency categories (from 0 = *not once* to 6 = *in general more than once a day*). The sum scores of the subscales range from 0-100, which were calculated as follows: (sum of item scores * 100)/ (number of items of the subscale * maximum item score). The sum score of the total NWFQ score was calculated based on the sum of all 47 items according to the same principle. A higher total score indicates more impairment in work functioning. Change scores of the NWFQ subscales were calculated as a T0-score minus T1-score; therefore, higher change scores indicate greater improvement.

Prior data of the same study population were used for the classification of subjects with high baseline scores (i.e., more impairments in work functioning).[10] The cut-off values were set at the 75th or 85th percentile of participants with mental health complaints in that study.

Anchor measure: important change in work functioning

As anchor for exploring the interpretability of change, one question assessed the participants' perceived important change in work functioning. Participants were asked the following question at the three-month follow-up: *"Compared to the last time you filled out this questionnaire (three months ago), do you feel your work functioning has changed importantly with regard to …?"* This question was answered for work functioning in general and six different aspects of work functioning in accordance with the six NWFQ subscales. A 5-point answering scale was used (*0 = yes, deteriorated; 1 = yes, deteriorated a little; 2 = no, not changed; 3 = yes, improved a little; 4 = yes, improved).* Participants were classified into three groups based on this scoring: 0 was classified as importantly deteriorated, 1, 2 and 3 as unchanged, and 4 as importantly improved.[15 21]

Data analysis

Descriptive factors were calculated for the demographic and occupational characteristics. To control for possible systematic loss to follow-up, we tested differences in socio demographic characteristics and work functioning impairments in the group of baseline respondents that were lost to follow-up and the group of respondents that filled out both questionnaires. A Mann-Whitney U test was performed for age in years, years of work experience, working hours per week, and for the separate NWFQ scales. A Fisher's exact test was performed for gender and ethnical background.

Anchor adequacy

The Spearman correlation coefficients between the anchor scores and the NWFQ change scores were calculated to evaluate anchor adequacy. The criterion for good anchor adequacy was set at > 0.30.[13]

Minimal Important Change (MIC)

We applied two different anchor-based methods for MIC calculation: 1) the mean-change method and 2) the ROC-curve method.

1) The mean-change method: The MIC was calculated using the mean change of the group who reported to perceive important improvement minus the mean change of the group who reported being unchanged.[15]

   MIC = MeanDiff improved - MeanDiff stable

2) ROC-curve method: The sensitivity and specificity in correctly distinguishing between unchanged and importantly improved subjects was calculated for each possible cut-off value of the NWFQ change scores. The anchor-based classification into importantly improved versus unchanged subjects was used as a reference standard, and the NWFQ change scores were used for the occurrence of minimally important change. The ROC-curve was constructed by plotting the sensitivity and the 1- specificity for all possible cut-off points. The MIC was defined as the cut-off value with the smallest number of misclassifications, i.e., the point at the upper left corner of the ROC-curve figure.[21 25]

Baseline correction for MIC calculation

Baseline scores are expected to impact MIC scores.[17 18 21 26] For example, small improvements in work functioning in subjects with high level of impairments are probably less meaningful than in subjects with low initial NWFQ scores. We applied two methods for baseline correction. First, separate MIC values were assessed for subjects with high baseline scores. No separate MIC values were assessed for the group of subjects with low baseline scores because improvement in this subgroup is barely possible. Second, MIC values were calculated based on the relative change scores, the percentage of change in relation to the baseline score [(T0-T1/T0)*100].

Smallest detectable change

The smallest detectable change (SDC) identifies the change scores that fall outside of the measurement error of a scale.[11] The calculation of the SDC was based on the standard error of measurement (SEM) corrected for the 95% criteria and the two measurement points. The SDC was represented by the formula:

$$SDC = 1.96 * \sqrt{2} * SEM$$

Subjects with change scores larger than the SDC have an at least 95% probability that a real change occurred, and the change score is not due to measurement error. However, subjects with change scores smaller than the SDC have a more than 5% chance that the change score was due to measurement error. SEM values from an earlier study, calculated from test-retest measures in 84 stable subjects of the same population as in the current study, were used.[10]

Evaluation of interpretability of change

We formulated two criteria for good interpretability of change: an area under the curve (AUC) for the ROC-curve method of 0.70 or higher[28] and MIC values, both mean-change and ROC based, larger than the SDC values.[11]

RESULTS

Of the 557 nurses and allied health professionals who participated in the RCT at baseline, 358 (64%) workers completed the second questionnaire. Table 1 presents the socio-demographic characteristics of the sample. The majority of the participants were female (80%) and native Dutch (87%). The mean age was 41 years, mean work experience was 12 years, and mean working hours was 30 per week. Two-thirds of the participants were nurses. No significant differences for social demographic characteristics and work functioning impairments were found between the group of respondents that filled out both baseline and follow-up questionnaires and the group of respondents that were lost to follow-up, except for the NWFQ subscale 'Impaired contact with patients and family ($p = 0.025$). The median in the dropout group was 13 whereas median in the group of completers was 8.

Table 2 presents the distribution of answers to the anchor question. Twenty-one (6%) subjects indicated to perceive important improvement in their work functioning on the overall NWFQ score, and 11 (3%) to 19 (5%) subjects on the six subscales. A very small number of subjects (1% maximum) indicated to perceive deterioration.

Mean change scores

The mean change scores of the total group revealed small improvements of 1 to 4 points on a 0 to 100 scale on all NWFQ subscales (Table 3). These improvements were larger in the high baseline group i.e., subjects with greater impairments in work functioning at baseline, with mean change scores of 5 to 12 points. The mean change scores for all scales were larger for the group that perceived important improvement than for the group that perceived to be unchanged or to be changed only a little. These differences were greater for the high baseline group.

Adequacy of the anchor

Table 4 presents the Spearman correlation coefficients of the NWFQ individual change scores with the individual scores on the anchor question as an indication for the adequacy of the anchor question. The correlations for the total group were low with only one correlation exceeding 0.30 and two correlations above 0.20. The correlations were higher in the high baseline group; all of the correlations were greater than 0.20, and three values were greater than 0.35. These differences in correlations indicate that subjects' perceived improvement in work functioning in the high baseline group was better represented by the NWFQ change scores than in the total group.

Table 1. Demographic characteristics of the participants (N = 358).

	N	(%)	Mean (SD)
Gender			
female	285	(80)	
Age in years			41 (12)
18-24	26	(7)	
25-34	100	(28)	
35-44	77	(22)	
45-54	98	(27)	
55-64	57	(16)	
Marital status			
Married/ living together with partner	245	(69)	
Partnered, not living together	17	(5)	
Single	79	(22)	
Divorced	10	(3)	
Widowed/widower	3	(1)	
Ethnic background			
Dutch	308	(87)	
Immigrant	48	(14)	
Occupation			
Nurse	227	(63)	
Nurse practitioner	34	(10)	
Surgical nurse	11	(3)	
Anesthetic nurse	11	(3)	
Allied health professional	57	(21)	
Nursing specialization (N = 227)			
Yes	144	(83)	
Years of experience			12 (10)
<=1	39	(11)	
2-5	98	(28)	
6-10	70	(20)	
11-15	52	(15)	
16-20	25	(7)	
21-25	20	(6)	
>25	50	(14)	
Working hours per week according to contract			
Type of contract			30 (7)
Permanent position	329	(93)	
Fixed-term contract	18	(5)	
Temporary employment	4	(1)	
Other	4	(1)	

Minimal Important Change

MIC values are presented in Table 4. The MIC based on the mean-change method ranged from 3.4 to 8.3 on the seven tested scales. The MIC based on the ROC-curve method ranged from 1.5 to 9.5. The range between these two methods was greatest (6.1 points) for the subscale *Impaired contact with patients and their family* and smallest (0.4 points) for the subscale *Causing incidents at work*. MIC values in the high baseline group were larger than in the total group with one exception for the mean-change method MIC of the subscale *Cognitive aspects of task execution and general errors*.

MIC values for the relative change scores ranged from 6% to 77% in the total group and from 8% to 76% in the high baseline group. Differences between the mean-change method and the ROC-curve method varied from 1% to 54% in the total group and 0% to 40% in the high baseline group.

Table 2. Frequencies of answers on the anchor questions for the total NWFQ score and the six subscales, at three months follow-up.

Perceived change	NWFQ total		1 cognitive		3 incidents		4 avoidance		5 colleagues		6 patients		7 energy & motivation	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
Total group	358	(100	358	(100)	249	(100)	358	(100)	358	(100)	358	(100)	358	(100)
Deteriorated														
Yes, deteriorated	1	(0)	2	(1)	0	(0)	0	(0)	0	(0)	0	(0)	5	(1)
Stable														
Yes, deteriorated a little	18	(5)	17	(5)	5	(2)	15	(4)	29	(8)	5	(1)	55	(15)
No, unchanged	276	(77)	277	(77	217	(87)	309	(86)	279	(78)	317	(89)	240	(67)
Yes, improved a little	42	(12)	43	(12)	16	(6)	18	(5)	37	(10)	25	(7)	43	(12)
Improved														
Yes, improved	21	(6)	19	(5)	11	(4)	16	(5)	13	(4)	11	(3)	15	(4)

Table 3. Median and range at baseline (T0) and three months follow-up (T1) and mean change scores and relative change scores for the total group (N = 358) and separately for the perceived unchanged and perceived improved group and also stratified for baseline values low versus high.

Scale	T0		T1		T0-T1 total group				T0-T1 perceived unchanged group				T0-T1 perceived improved group			
	Median	(range)	Median	(range)	N	Mean	(SD)	Mean rel. change	N	Mean	(SD)	Mean rel. change	N	Mean	(SD)	Mean rel. change
Total group																
NWFQ total scores	11	(0-51)	8	(0-39)	358	2	(6)	9%	336	2	(6)	9%	21	6	(7)	15%
1. Cognitive aspects of task execution	15	(0-73)	14	(0-65)	358	4	(12)	1%	337	3	(11)	-1%	19	12	(10)	39%
3. Causing incidents at work	6	(0-42)	4	(0-44)	249	2	(5)	13%	238	1	(5)	12%	11	6	(8)	48%
4. Avoidance behavior	3	(0-59)	3	(0-53)	358	1	(7)	1%	342	0	(7)	0%	16	6	(12)	22%
5. Conflicts and annoyances with colleagues	7	(0-93)	4	(0-68)	358	2	(11)	7%	345	2	(11)	5%	13	6	(15)	51%
6. Impaired contact with patients and family	8	(0-56)	7	(0-46)	358	3	(9)	9%	347	3	(9)	8%	11	6	(9)	31%
7. Lack of energy and motivation	7	(0-73)	7	(0-87)	358	1	(11)	-8%	338	1	(9)	-3%	15	8	(14)	24%
Low baseline group																
NWFQ total scores	9	(0-21)	7	(0-38)	305	1	(5)	7%	289	1	(5)	7%	15	4	(5)	9%
1. Cognitive aspects of task execution	11	(0-24)	11	(0-45)	257	1	(9)	-11%	247	1	(9)	-12%	9	5	(8)	29%
3. Causing incidents at work	4	(0-13)	2	(0-19)	202	1	(4)	11%	193	1	(4)	10%	9	3	(3)	40%
4. Avoidance behavior	3	(0-9)	0	(0-44)	273	-1	(6)	-7%	265	-1	(6)	-7%	8	0	(8)	1%
5. Conflicts and annoyances with colleagues	7	(0-25)	4	(0-61)	315	1	(9)	4%	305	1	(9)	2%	10	2	(9)	57%
6. Impaired contact with patients and family	8	(0-17)	4	(0-29)	276	1	(6)	2%	267	1	(6)	1%	9	3	(6)	20%
7. Lack of energy and motivation	7	(0-30)	3	(0-87)	323	0	(9)	-12%	305	0	(8)	-6%	13	4	(7)	18%
High baseline group																
NWFQ total scores	28	(22-51)	23	(1-39)	53	7	(9)	24%	47	7	(8)	23%	6	11	(10)	31%
1. Cognitive aspects of task execution	36	(26-73)	26	(0-65)	101	12	(14)	31%	90	12	(14)	29%	10	18	(9)	48%
3. Causing incidents at work	19	(15-42)	15	(0-44)	47	5	(9)	23%	45	5	(8)	21%	2	22	(1)	81%
4. Avoidance behavior	19	(13-59)	13	(0-53)	85	5	(10)	27%	77	4	(9)	25%	8	12	(14)	44%
5. Conflicts and annoyances with colleagues	39	(29-93)	32	(0-68)	43	12	(17)	27%	40	11	(17)	27%	3	20	(24)	32%
6. Impaired contact with patients and family	23	(19-56)	17	(0-46)	82	9	(13)	32%	80	9	(13)	30%	2	20	(1)	80%
7. Lack of energy and motivation	43	(33-73)	30	(7-67)	35	11	(16)	25%	33	9	(14)	22%	2	39	(12)	62%

Mean rel. Change = mean relative change scores: (T0-T1/T0) * 100

Table 4. Correlations of change scores and the anchor, SEM, SDC, and MIC scores for improvement for the total NWFQ score and the six NWFQ subscales, for the total group and the high baseline group.

	NWFQ total		1 cognitive		3 incidents		4 avoidance		5 colleagues		6 patients		7 energy & motivation	
Spearman correlations: anchor with change score														
Total group	0.13		0.19		0.13		0.07		0.22		0.05		0.30	
Low baseline group	0.08		0.10		0.10		0.01		0.13		-0.05		0.30	
High baseline group	0.26		0.25		**0.36**		0.21		**0.45**		0.24		**0.48**	
SEM*	2.61		4.27		3.04		3.33		6.03		2.95		6.12	
SDC	7.2		11.8		8.4		9.2		16.7		8.2		17	
MIC														
Total group														
MIC method 1: mean change	3.9		8.3		5.1		5.4		4.2		3.4		7.4	
MIC method 2: ROC curve (AUC)	1.5	(0.66)	6.5	**(0.72)**	5.5	**(0.70)**	1.5	(0.66)	1.5	(0.58)	**9.5**	(0.58)	6.5	(0.64)
MIC method 1: mean relative change	6%		40%		36%		22%		46%		23%		27%	
MIC method 2: ROC curve relative change (AUC)	24%	(0.65)	23%	(0.64)	47%	(0.68)	12%	(0.60)	45%	(0.63)	77%	(0.55)	23%	(0.58)
High baseline group														
MIC method 1: mean change	4.4		6.3		**17**		7.6		9.2		**11.3**		**29**	
MIC method 2: ROC curve (AUC)	**15**	(0.63)	**13.5**	(0.64)	**19**	**(0.98)**	**9.5**	(0.65)	**41.5**	(0.63)	**18.5**	**(0.85)**	**28.5**	**(0.98)**
MIC method 1: mean relative change	8%		19%		61%		19%		4%		50%		40%	
MIC method 2: ROC curve relative change (AUC)	40%	(0.57)	23%	(0.66)	61%	**(0.93)**	12%	(0.60)	44%	(0.51)	76%	**(0.86)**	55%	**(0.86)**

Bold printed value: exceeds the minimum value for the criterion for good interpretability of change.
* results based on a previous study

Table 5. Number of subjects with change score equal to or higher than the SDC.

NWFQ Scale	SDC	Change scores above SDC								
		Total group			Perceived unchanged group			Perceived improved group		
		N	N (T0-T1 > SDC)	(%)	N	N (T0-T1 > SDC)	(%)	N	N (T0-T1 > SDC)	(%)
Total group										
NWFQ total scores	7.2	358	59	(17)	336	52	(16)	21	7	(33)
1. Cognitive aspects of task execution	11.8	358	84	(24)	337	74	(22)	19	10	(53)
3. Causing incidents at work	8.4	249	19	(8)	238	16	(7)	11	3	(27)
4. Avoidance behavior	9.2	358	31	(9)	342	6	(80	16	5	(31)
5. Conflicts and annoyances with colleagues	16.7	358	29	(8)	345	27	(8)	13	2	(15)
6. Impaired contact with patients and family	8.2	358	68	(19)	347	64	(18)	11	4	(36)
7. Lack of energy and motivation	17	358	17	(5)	338	15	(4)	15	2	(13)
High baseline group										
NWFQ total scores	7.2	23	25	(47)	47	21	(45)	6	4	(67)
1. Cognitive aspects of task execution	11.8	26	56	(55)	90	48	(53)	10	8	(80)
3. Causing incidents at work	8.4	15	15	(32)	45	13	(29)	2	2	(100)
4. Avoidance behavior	9.2	13	31	(37)	77	26	(34)	8	5	(63)
5. Conflicts and annoyances with colleagues	16.7	32	21	(49)	40	19	(48)	3	2	(67)
6. Impaired contact with patients and family	8.2	17	44	(54)	80	42	(53)	2	2	(100)
7. Lack of energy and motivation	17	30	12	(34)	33	10	(30)	2	2	(100)

The Smallest Detectable Change

The Smallest Detectable Change (SDC) for the total NWFQ scale was 7.2 (Table 4). The SDCs for the six subscales ranged from 8.2 to 11.8 for four subscales, and two subscales exhibited higher SDC values (16.7 and 17). As presented in Table 5, 32% to 55% of the change scores in the high baseline group show a positive change score that exceeded the SDC of the NWFQ scales. The subjects in the high baseline group who perceived an important improvement show higher percentages of change scores (from 63% to 100%) that were equal to or higher than the SDC.

The interpretability of change in the NWFQ

AUC values

The AUC values based on the absolute change scores for the total sample ranged from 0.58 to 0.72 (see Table 4). Two AUC values met the criterion of ≥ 0.70. The AUC values in the high baseline group were higher than in the total group, ranging from 0.63 to 0.98. The AUC values of three subscales met the ≥ 0.70 criterion.

The AUC values based on relative change scores were all slightly lower, ranging from 0.55 to 0.68 in the total group, none exceeding the 0.70 limit. The AUC values based on relative change scores in the high baseline group ranged from 0.51 to 0.98, three scale values exceed the 0.70 criterion.

MIC values in relation to SDC values

Only one of the 14 MIC values for the total group exceeded the SDC value, the ROC-curve based MIC (9.5) for *Impaired contact with patients and their family.* Ten of the 14 values exceeded the SDC in the high baseline group. Here, all ROC-curve based MIC values (from 9.5 to 41.5) and three mean-change MIC values exceeded the SDC values: *Causing incidents at work* (MIC = 17), *Impaired contact with patients and their family* (MIC = 11.3), *and Lack of energy and motivation* (MIC = 29), which suggests a good interpretability of change in the high baseline group.

DISCUSSION

The purpose of this study was to evaluate the interpretability of individual changes identifying the MIC and the SDC of the overall NWFQ score and the six subscales in the preventive setting of a workers' health surveillance in nurses and allied health professionals.

The requirements for a good interpretability of change were not met for the total group. However, three of the seven tested scales in the high baseline group fulfilled the requirements: *Causing incidents at work, Impaired contact with patients and their family,* and

Lack of energy and motivation. The SDC values were small compared to the scale range, except for two values. SDCs ranged from 7.2 to 17.

Theoretical frame for interpretation of results

We conclude that evaluating the ability to measure improvements using the NWFQ in a preventive setting is difficult because only a small proportion of the population perceived improved work functioning. Therefore, we chose to focus on the high baseline group results. This choice was based on the population in which the data is collected and on the characteristics of the tested scale. The current RCT in which the data for this study was collected was conducted in a working population of primarily healthy workers who functioned well at work overall. Only a limited number of workers scored above the cut-off for impairments in work functioning at baseline (between 10% and 28%) because of this preventive setting. Improvement in workers with no or limited impairments in work functioning at baseline cannot be captured by NWFQ change scores because the NWFQ is constructed to measure impaired work functioning. This suggests that improvement on the NWFQ subscales was unlikely in approximately three quarters of the sample. As improvement in the group of subjects with low NWFQ scores is not realistic, MIC values for improvement for the total group are less relevant than for the high baseline group.

The used anchor and its adequacy

The correlations between the anchor score and the NWFQ change score were low in the total group (from 0.07 to 0.30). As pointed out above, three quarters of the sample were unlikely to improve on the NWFQ scales; however, these subjects were able to indicate improvement on the anchor question. In other words, improvements assessed by the anchor in subjects with good work functioning at baseline may not be adequately reflected in the NWFQ change scores on work functioning impairments, which makes these low correlations rational. Therefore the anchor is not adequate for the use in the total group.

The correlations in the high baseline group were higher than in the total group, three of the seven correlations for the high baseline group met the required 0.30.[13] These correlations were similar to and better than the correlations found by Beaton and colleagues, who evaluated the responsiveness of five work functioning questionnaires in arthritis patients.[28] Although the current results are better than a comparable study, the validity of the anchor quality remains questionable. A number of aspects might affect the validity of the anchor-question, e.g., recall bias.[17] Unfortunately, we have no information on the validity and reliability of the used anchor other than the presented correlations. The low correlation coefficients for four of the investigated NWFQ scales and the rather low correlations in the study of Beaton for five different work functioning instruments, let us assume that formulating a valid single item anchor score on work functioning is complex and might be a too simple representation of what is thought to be measured in relation to changed work functioning. Therefore, further

validation of single item anchor questions on work functioning are needed as well as more insight into the role of recall bias and the measurement precision.

What do the partly insufficient correlations between anchor and actual change score mean for the application of the found MIC values? Terwee and colleagues concluded that correlations between anchor and scale changes do not directly influence the MIC value itself.[21] Therefore, we regard the correlations sufficient for interpretations of the high baseline group MIC values. Still, low associations between NWFQ change scores and anchor affect the interpretability because they lead to less favorable AUC values.

Application of the found NWFQ MIC and SDC values in research and practice

Three of the tested scales met the requirements for a good interpretability of change. However, we take into account that our results were based on a small number of subjects, and that the MIC is only applicable to subjects that score high at baseline. We consider for the remaining four scales that it is too early to state that change scores cannot be interpreted, given the doubt about the anchor question used and the small N. De Vet and colleagues argue that not meeting the requirement that the MIC exceeds the SDC is no reason to state that the change scores of these scales cannot be interpreted. In practice the awareness of the larger possibility of false positive classifications should be promoted.[20] We conclude that although the observed MIC values can be applied in case of high baseline value, it has to be done with caution. Future research should further evaluate the interpretability of change scores.

This study produced various MIC values for the NWFQ scales due to the application of two anchor-based MIC calculation methods and two methods for baseline correction. Some of these values differed from each other with a relatively high range. We recommend using only one MIC value when applying the MIC values of this study for the interpretation of an individual's change score in practice or in research. One might either use the mean of both MIC values (mean-change and ROC-curve method) or chose one of the two MIC values, either the highest or the lowest, depending on the preference for a reduced risk for false positives or false negatives.

Concerning the different baseline correction methods, we conclude that relative change score MIC of the total group do not sufficiently correct for baseline differences in this setting. The differences in relative change score MIC between the total and high baseline group support this conclusion. We recommend the use of the relative change MIC values of the high baseline group to best allow for baseline differences in research and practice. However, using relative change scores MICs by practitioners is complex as it asks for additional calculations.

To illustrate the application of SDC and MIC values, we present an example for the NWFQ total score. Here, we use absolute change score MICs for more simplicity. A practitioner can conclude that improvements of seven points or more on the NWFQ total score in high baseline value workers can be regarded "real" changes, with 95% confidence (Table 5). However, a

practitioner does not know if these improvements are also of relevance, without applying the MIC value. The mean of the two MIC values of absolute change scores is 10 (Table 4). Therefore, the practitioner can conclude that any improvements higher than 10 points can be regarded "real" and relevant changes in individual high baseline workers.

The 95% criteria applied for the calculation of the SDC might be perceived as very strict by professionals in occupational health practice, as they might find lower confidence for a real change sufficient, i.e., 60% confidence. In that case a professional might apply lower SDC values.

The SDC scores for the NWFQ scales are related to change scores of individuals. When applying SDC on group level, e.g., for the interpretation of differences between an intervention and a control group, one could correct for the number of subjects by dividing the SDC with the square root of the number of subjects in that study.[29 30]

Methodological notes

The most distinct limitation of this study is the low number of subjects on which a number of calculations were based. Two methodologies impacted this number: first, the anchor classification and second, the chosen cut-off points for the high baseline group identification.

We chose, consistent with previous studies,[21 31] to classify the outer category of the original 5-point anchor-question as "improved", and the "improved a little", and "deteriorated a little" as unchanged. The subjects in our sample were primarily healthy, well-functioning workers who participated at work. We assumed that healthy subjects who reported to perceive "a little improved" work functioning indicated no important changes in this prevention context. Turner and colleagues[32] recommended the classification of the category adjacent to the category "unchanged" as improved rather than unchanged in a study of different classification methods for anchor-based MIC calculations. They concluded that this classification would lead to larger AUC values and narrower confidence intervals, which increased confidence in the estimated MIC points. However, post-hoc analyses revealed that this was not true for our data (results not presented), comparing the AUC values of both anchor classification methods. These results support our choice of anchor classification.

The analyses of the high baseline subgroup were based on a number of subjects below 50 for three scales, leading to low numbers of subjects that perceived improvement. The cut-off values applied were the same as the values for case identification during the screening part of the workers' health surveillance mental module. We explored the use of a lower cut-off value for the identification of subjects with impairments in work functioning for the selection of a high baseline group running post-hoc analyses of interpretability and MIC for the new high baseline group (supplemental file Table S1 and S2). The N for the high baseline workers rises, but the N of subjects that perceived improvement only slightly improved. It is noticeable however, that the AUC values only differed slightly from the original high baseline group

despite this larger N. MIC values based on the relative change score barely differed between the original and the new analysis with a lower cut-off, in contrast to the MIC values that were based on absolute change scores.

Overall, we conclude that the methodological choices in the study did not negatively influence the observed results.

Recommendations for further research

Suitable anchor questions should be identified in questionnaires on work functioning to guide the calculation of MIC values. The results of the present study and of a study on work functioning questionnaires by Beaton and colleagues [28] demonstrated that the use of a single item self-report anchor exhibited only a small overlap with change scores on the measurement instruments. The formulation of an accurate anchor question for perceived change in work functioning is complex. A study using verbal probe interview techniques might be useful to understand the process of anchor question interpretation and answering, which might help identify a more suitable formulation.[33]

The assessment of the MIC for deterioration is recommended in future studies, especially for the use in workers' health surveillance administered for monitoring purposes.

Furthermore, although no sample size indications for MIC calculation are provided in the literature yet.[21] we recommend the replication of the MIC calculation in a population in which the occurrence of important improvements in work functioning is expected to be higher.

REFERENCES

1 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
2 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
3 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
4 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
5 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
6 Hees HL, Koeter MW, de Vries G et al. Effectiveness of adjuvant occupational therapy in employees with depression: design of a randomized controlled trial. *BMC Public Health* 2010;10:558.
7 Abma FI, van der Klink JJ, Terwee CB et al. Evaluation of the measurement properties of self-reported health-related work-functioning instruments among workers with common mental disorders. *Scand J Work Environ Health* 2012;38:5-18.
8 Nieuwenhuijsen K, Franche RL, van Dijk FJ. Work functioning measurement: tools for occupational mental health research. *J Occup Environ Med* 2010;52:778-90.
9 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. Impaired work functioning due to common mental disorders in nurses and allied health professionals: the Nurses Work Functioning Questionnaire. *Int Arch Occup Environ Health* 2011;85:125-38.
10 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. Psychometric Properties of the Nurses Work Functioning Questionnaire (NWFQ). *PLoS One* 2011;6:e26565.
11 de Vet HC, Terwee CB, Ostelo RW et al. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes* 2006;4:54.
12 Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002;14:109-14.
13 Revicki D, Hays RD, Cella D et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102-9.
14 Beckerman H, Roebroeck ME, Lankhorst GJ et al. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* 2001;10:571-8.
15 Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407-15.
16 Beaton DE, Bombardier C, Katz JN et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. *J Rheumatol* 2001;28:400-5.
17 Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395-407.
18 Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. *J Clin Epidemiol* 2004;57:1153-60.
19 de Vet HC, Ostelo RW, Terwee CB et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007;16:131-42.
20 de Vet HC, Terluin B, Knol DL et al. Three ways to quantify uncertainty in individually applied "minimally important change" values. *J Clin Epidemiol* 2010;63:37-45.
21 Terwee CB, Roorda LD, Dekker J et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;63:524-34.
22 Hageman WJ, Arrindell WA. Establishing clinically significant change: increment of precision and the distinction between individual and group level of analysis. *Behav Res Ther* 1999;37:1169-93.
23 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582-92.
24 Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999;52:861-73.
25 Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986;39:897-906.
26 Kolotkin RL, Crosby RD. Psychometric evaluation of the impact of weight on quality of life-lite questionnaire (IWQOL-lite) in a community sample. *Qual Life Res* 2002;11:157-71.

27 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.
28 Beaton DE, Tang K, Gignac MA et al. Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. *Arthritis Care Res (Hoboken)* 2010;62:28-37.
29 de Boer MR, de Vet HC, Terwee CB et al. Changes to the subscales of two vision-related quality of life questionnaires are proposed. *J Clin Epidemiol* 2005;58:1260-8.
30 de Vet HC, Bouter LM, Bezemer PD et al. Reproducibility and responsiveness of evaluative outcome measures. Theoretical considerations illustrated by an empirical example. *Int J Technol Assess Health Care* 2001;17:479-87.
31 Hendriks EJ, Bernards AT, Berghmans BC et al. The psychometric properties of the PRAFAB-questionnaire: a brief assessment questionnaire to evaluate severity of urinary incontinence in women. *Neurourol Urodyn* 2007;26:998-1007.
32 Turner D, Schunemann HJ, Griffith LE et al. Using the entire cohort in the receiver operating characteristic analysis maximizes precision of the minimal important difference. *J Clin Epidemiol* 2009;62:374-9.
33 Willis GB. Cognitive interviewing in practice: think-aloud, verbal probing and other techniques. Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks, CA: Sage Publications, 2005: 42-63.

SUPPLEMENT

Table S1. Median and range at baseline (T0) and three months follow-up (T1) and mean change scores and relative change scores for the total group (N = 358) and separately for the perceived unchanged and perceived improved group and also stratified for baseline values low versus high.

Scale	T0		T1		T0-T1 total group				T0-T1 perceived unchanged group				T0-T1 perceived improved group			
	Median	(range)	Median	(range)	N	Mean	(SD)	Mean rel. change	N	Mean	(SD)	Mean rel. change	N	Mean	(SD)	Mean rel. change
Total group																
NWFQ total scores	11	(0-51)	8	(0-39)	358	2	(6)	9%	336	2	(6)	9%	21	6	(7)	15%
1. Cognitive aspects of task execution	15	(0-73)	14	(0-65)	358	4	(12)	1%	337	3	(11)	-1%	19	12	(10)	39%
3. Causing incidents at work	6	(0-42)	4	(0-44)	249	2	(5)	13%	238	1	(5)	12%	11	6	(8)	48%
4. Avoidance behavior	3	(0-59)	3	(0-53)	358	1	(7)	1%	342	0	(7)	0%	16	6	(12)	22%
5. Conflicts and annoyances with colleagues	7	(0-93)	4	(0-68)	358	2	(11)	7%	345	2	(11)	5%	13	6	(15)	51%
6. Impaired contact with patients and family	8	(0-56)	7	(0-46)	358	3	(9)	9%	347	3	(9)	8%	11	6	(9)	31%
7. Lack of energy and motivation	7	(0-73)	7	(0-87)	358	1	(11)	-8%	338	1	(9)	-3%	15	8	(14)	24%
Low baseline group																
NWFQ total scores	9	(0-21)	7	(0-38)	305	1	(5)	7%	289	1	(5)	7%	15	4	(5)	9%
1. Cognitive aspects of task execution	11	(0-24)	11	(0-45)	257	1	(9)	-11%	247	1	(9)	-12%	9	5	(8)	29%
3. Causing incidents at work	4	(0-13)	2	(0-19)	202	1	(4)	11%	193	1	(4)	10%	9	3	(3)	40%
4. Avoidance behavior	3	(0-9)	0	(0-44)	273	-1	(6)	-7%	265	-1	(6)	-7%	8	0	(8)	1%
5. Conflicts and annoyances with colleagues	7	(0-25)	4	(0-61)	315	1	(9)	4%	305	1	(9)	2%	10	2	(9)	57%
6. Impaired contact with patients and family	8	(0-17)	4	(0-29)	276	1	(6)	2%	267	1	(6)	1%	9	3	(6)	20%
7. Lack of energy and motivation	7	(0-30)	3	(0-87)	323	0	(9)	-12%	305	0	(8)	-6%	13	4	(7)	18%

Mean rel. change = mean relative change scores: (T0-T1/T0) * 100

Table S1. (continued)

Scale	T0		T1		T0-T1 total group				T0-T1 perceived unchanged group				T0-T1 perceived improved group			
	Median (range)		Median (range)		N	Mean (SD)		Mean rel. change	N	Mean (SD)		Mean rel. change	N	Mean (SD)		Mean rel. change
High baseline group (original cutoff)																
NWFQ total scores	28	(22-51)	23	(1-39)	53	7	(9)	24%	47	7	(8)	23%	6	11	(10)	31%
1. Cognitive aspects of task execution	36	(26-73)	26	(0-65)	101	12	(14)	31%	90	12	(14)	29%	10	18	(9)	48%
3. Causing incidents at work	19	(15-42)	15	(0-44)	47	5	(9)	23%	45	5	(8)	21%	2	22	(1)	81%
4. Avoidance behavior	19	(13-59)	13	(0-53)	85	5	(10)	27%	77	4	(9)	25%	8	12	(14)	44%
5. Conflicts and annoyances with colleagues	39	(29-93)	32	(0-68)	43	12	(17)	27%	40	11	(17)	27%	3	20	(24)	32%
6. Impaired contact with patients and family	23	(19-56)	17	(0-46)	82	9	(13)	32%	80	9	(13)	30%	2	20	(1)	80%
7. Lack of energy and motivation	43	(33-73)	30	(7-67)	35	11	(16)	25%	33	9	(14)	22%	2	39	(12)	62%
High baseline group (new, lower cut-off)																
NWFQ total scores	25	(20-51)	21	(1-39)	72	6	(8)	23%	65	6	(8)	22%	7	11	(10)	32%
1. Cognitive aspects of task execution	35	(23-73)	23	(0-65)	117	11	(13)	30%	104	11	(14)	28%	12	18	(8)	51%
3. Causing incidents at work	17	(13-42)	13	(0-44)	61	5	(80	24%	58	4	(80	21%	3	17	(7)	77%
4. Avoidance behavior	19	(13-59)	13	(0-53)	85	5	(10)	27%	77	4	(9)	25%	8	12	(14)	44%
5. Conflicts and annoyances with colleagues	27	(18-93)	20	(0-68)	86	8	(16)	23%	83	7	(16)	23%	3	20	(24)	32%
6. Impaired contact with patients and family	21	(15-56)	13	(0-46)	121	8	(11)	34%	116	8	(11)	33%	5	12	(9)	61%
7. Lack of energy and motivation	37	(27-73)	29	(7-87)	52	7	(18)	46%	46	8	(14)	20%	4	27	(15)	58%

Mean rel. change = mean relative change scores: (T0-T1/T0) * 100

Table S2. Correlations of change scores and the anchor, SEM, SDC, and MIC scores for improvement for the total NWFQ score and the different subscales for the total group and the high baseline group.

	NWFQ total		1 cognitive		3 incidents		4 avoidance		5 colleagues		6 patients		7 energy & motivation	
Spearman correlations: anchor with change score														
Total group	0.13		0.19		0.13		0.07		0.22		0.05		0.30	
Low baseline group	0.08		0.10		0.10		0.01		0.13		-0.05		0.30	
High baseline group **(original cut-off)**	0.26		0.25		**0.36**		0.21		**0.45**		0.24		**0.48**	
High baseline group **(new, lower cut-off)**	0.23		0.21		**0.30**		0.21		**0.40**		0.18		**0.51**	
SEM*	2.61		4.27		3.04		3.33		6.03		2.95		6.12	
SDC	7.2		11.8		8.4		9.2		16.7		8.2		17	
MIC														
Total group														
MIC method 1: mean change	3.9		8.3		5.1		5.4		4.2		3.4		7.4	
MIC method 2: ROC curve (AUC)	1.5	(0.66)	6.5	**(0.72)**	5.5	**(0.70)**	1.5	(0.66)	1.5	(0.58)	**9.5**	(0.58)	6.5	(0.64)
MIC method 1: mean relative change	6%		40%		36%		22%		46%		23%		27%	
MIC method 2: ROC curve relative change (AUC)	24%	(0.65)	23%	(0.64)	47%	(0.68)	12%	(0.60)	45%	(0.63)	77%	(0.55)	23%	(0.58)
High baseline group (original cut-off)														
MIC method 1: mean change	4.4		6.3		**17**		7.6		9.2		**11.3**		**29**	
MIC method 2: ROC curve (AUC)	**15**	(0.63)	**13.5**	(0.64)	**19**	**(0.98)**	**9.5**	(0.65)	**41.5**	(0.63)	**18.5**	**(0.85)**	**28.5**	**(0.98)**
MIC method 1: mean relative change	8%		19%		61%		19%		4%		50%		40%	
MIC method 2: ROC curve relative change (AUC)	40%	(0.57)	23%	(0.66)	61%	**(0.93)**	12%	(0.60)	44%	(0.51)	76%	**(0.86)**	55%	**(0.86)**
High baseline group (new, lower cut-off)														
MIC method 1: mean change	4.6		6.9		**13.2**		7.6		13		4.4		**18.9**	
MIC method 2: ROC curve (AUC)	6.5	(0.64)	9.5	(0.66)	**8.5**	**(0.91)**	**9.5**	(0.65)	**21.5**	(0.69)	**14.5**	**(0.65)**	11.5	**(0.82)**
MIC method 1: mean relative change	9%		23%		57%		19%		9%		28%		38%	
MIC method 2: ROC curve relative change (AUC)	35%	(0.59)	23%	(0.69)	61%	**(0.91)**	12%	(0.60)	44%	(0.52)	77%	**(0.69)**	41	**(0.83)**

Bold printed value: exceeds the minimum value for the criterion for good interpretability of change.
* results based on a previous study

6

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

Fania R. Gärtner
Sarah M. Ketelaar
Odile Smeets
Linda Bolier
Eva Fischer
Frank J.H. van Dijk
Karen Nieuwenhuijsen
Judith K. Sluiter

BMC Public Health 2011;11:290.

ABSTRACT

Background: Employees in health care service are at high risk for developing mental health complaints. The effects of mental health complaints on work can have serious consequences for the quality of care provided by these workers. To help health service workers remain healthy and productive, preventive actions are necessary. A Workers' Health Surveillance (WHS) mental module may be an effective strategy to monitor and promote good (mental) health and work performance. The objective of this paper is to describe the design of a three arm cluster randomized controlled trial on the effectiveness of a WHS mental module for nurses and allied health professionals. Two strategies for this WHS mental module will be compared along with data from a control group. Additionally, the cost effectiveness of the approaches will be evaluated from a societal perspective.

Methods: The study is designed as a cluster randomized controlled trial consisting of three arms (two intervention groups, 1 control group) with randomization at ward level. The study population consists of 86 departments in one Dutch academic medical center with a total of 1731 nurses and allied health professionals. At baseline, after three months and after six months of follow-up, outcomes will be assessed by online questionnaires. In both intervention arms, participants will complete a screening to detect problems in mental health and work functioning and receive feedback on their screening results. In cases of impairments in mental health or work functioning in the first intervention arm, a consultation with an occupational physician will be offered. The second intervention arm offers a choice of self-help e-mental health interventions, which will be tailored based on each individual's mental health state and work functioning. The primary outcomes will be help-seeking behavior and work functioning. Secondary outcomes will be mental health and wellbeing. Furthermore, cost-effectiveness in both intervention arms will be assessed, and a process evaluation will be performed.

Discussion: When it is proven effective compared to a control group, a WHS mental module for nurses and allied health professionals could be implemented and used on a regular basis by occupational health services in hospitals to improve employees' mental health and work functioning.

Trial Registration: NTR2786

BACKGROUND

Common mental disorders (CMDs) can have negative effects on work as they can impair work functioning and increase sickness absence.[1-5] In some occupations, the impairments in work functioning can have serious consequences, such as injuries to workers. One occupation in which this vulnerability is highly present is nursing. Nurses, the largest occupational group in healthcare, are at higher risk of developing mental health problems compared to workers outside of the health care sector and compared to other (health) service workers.[6] The relative risk for depression is high for nurses, RR = 3.5 (95% CI: 1.3 – 9.6), compared to other human service workers and other healthcare workers.[6] This high risk might partly be explained by the very nature of the work, with work environment characteristics that are known to promote the occurrence of mental health complaints, such as high job demands, low job control and low social support.[7 8] Furthermore, in the health care sector, impairments in work functioning can have serious effects not only for the nurses but also for patients and their safety as a recent literature review showed.[8 9]

In the Netherlands, the occupational health care that is provided for employees with mental health problems can be considered effective. Care according to the guidelines for occupational physicians' (OP) treatment of workers with mental health problems has been proven to improve mental health and to enhance return to work for sick-listed employees.[10 11] However, the health service for OPs is often not used by workers until they are sick-listed. Late or no help-seeking for mental health complaints is a well-known problem inside and outside of the occupational health servic.[12 13] Preventive actions are needed to provide timely help before work functioning is reduced to the extent that workers cause serious incidents or must call in sick. Early identification of health complaints and risks in work functioning to provide timely help is a first step in the prevention of more serious consequences for the health and safety of the nurses and their patients. Furthermore, preventive actions can improve the wellbeing of employees in the health care sector. Wellbeing can have positive effects on the engagement and productivity of employees. With the age of the caring workforce increasing, the importance of sustainable labor participation by senior employees is increasing. Therefore, it is of utmost importance to keep the caring work force engaged and mentally fit so they can continue to meet the high mental demands of modern-day work.

A Workers' Health Surveillance (WHS) mental module may be a successful preventive strategy for CMDs and impairments in work functioning in the health care sector. Within the occupational health care setting, WHS is a well-developed strategy for preventive actions.[14 15] WHS aims to detect negative health effects of work in an early stage to enable timely interventions.[15] Although the use and application of WHS is rising for various occupations and health effects, little is known about WHS targeting mental health effects. In a recent literature review by Plat et al.[16] on WHS in military and emergency service personnel, three studies included psychological health aspects, one in police personnel,[17] one in rescue and recovery

workers[18] and one in soldiers.[19] WHS for mental health effects in nurses has not yet been scientifically evaluated. Therefore, the aim of this study is to test the effectiveness of a job-specific WHS mental module for nurses and allied health professionals.

Although the International Labour Organization has formulated recommendations for the use of WHS, the design differs between countries. In the Netherlands, a policy guideline on how to conduct WHS exists.[20] This guideline does not prescribe any specific interventions, but includes principles and leading criteria such as the statement that screening for health problems should only be conducted if effective interventions for that health problem are available. Furthermore, one of the core aims of the guideline is the monitoring and improvement of both the health and functioning of workers. Therefore, our job-specific WHS mental module includes screening for early signals of mental health complaints and for impairments in work functioning. For the detection of impaired mental health, several validated instruments exist that are suited for the working population. However, until recently, no instrument for detecting impaired work functioning in healthcare workers related to mental health problems was available. Such an instrument has now been developed to be used in the hospital environment, the Nurses Work Functioning Questionnaire (NWFQ) (Gärtner, Nieuwenhuijsen, van Dijk, Sluiter, unpublished). The NWFQ was designed based on literature studies and focus group investigations with the workers' supervisors, human resource managers and occupational health professionals. The NWFQ has a high content validity, and its seven subscales show good or acceptable internal consistency.

For the interventions that follow the screening, two different strategies were developed. The first strategy is a consultation offered by the OP following a protocol for care for workers with mental health complaints, as developed for this study. The second strategy is a choice of self-help e-mental health interventions that is offered to all workers – those with and without complaints. The choice is tailored to the individual screening results.

The objective of the Mental Vitality @ Work study is to study the effectiveness of two strategies for the WHS mental module in one cluster randomized controlled trial design with three arms. Substudy 1 aims to test the effectiveness of screening for problems in mental health and work functioning plus advice on appropriate interventions by an OP compared to a control group. It will study the effects on adequate help-seeking behavior, work functioning and mental health. We hypothesize that employees who receive the WHS mental module with screening plus an invitation for OP-care will show more adequate help-seeking behavior than employees in the control group. Furthermore, we hypothesize that work functioning and mental health problems will improve in employees who receive the WHS mental module with screening plus invitation for OP-care compared to employees in the control group.

Substudy 2 aims to compare the OP-care strategy with a second strategy, including the same screening of problems in mental health and work functioning as in the OP-care strategy plus a stepped care e-mental health approach. Substudy 2 will compare the effects of both

strategies on work functioning and mental health. We hypothesize that both WHS mental module strategies are equivalent in their effectiveness on work functioning, mental health and wellbeing compared to the control group. An economic evaluation of the WHS mental module will be conducted alongside the randomized controlled trial. Cost-effectiveness of the WHS mental module will be assessed from a societal perspective. The employer's perspective will be considered in a cost-benefit analysis. Regarding cost effectiveness, we hypothesize that the WHS mental module with E-mental health interventions is more cost effective than the WHS mental module with OP-care.

METHODS/DESIGN

In the following description of the design of the Mental Vitality @ Work study we follow the CONSORT statement, which aims to improve the quality of reporting randomized controlled trials (RCT).[21 22]

Study design

A cluster randomized controlled trial with three parallel arms will be performed in order to evaluate the effectiveness of two strategies for a WHS mental module for nurses and allied health professionals: the OP-care strategy and the E-mental health strategy. The study combines two separately funded substudies. Substudy 1 will compare the control arm with the OP-care arm – a screening on mental health complaints and impaired work functioning followed by a consultation with an OP and appropriate interventions if necessary. Substudy 1 will test the effect on help-seeking behavior, work functioning and mental health of employees with problems in mental health and/or work functioning.

Substudy 2 will compare the E-mental health arm – a screening on mental health complaints and impaired work functioning followed by a tailored choice of self-help e-mental health interventions – with the control arm and with the OP-care arm. Additionally, a subgroup analysis of the healthy participants comparing the E-mental health arm with the OP-care arm will be conducted. Substudy 2 will test the effect of the interventions on work functioning and mental health.

Participants will be followed for six months. Two follow-up measures will be conducted, one at three months and one at six months. The Medical Ethics Committee of the Academic Medical Center in Amsterdam (AMC) gave approval for the study. Figure 1 presents an overview of the study design. Below, differences between Substudy 1 and Substudy 2 are described. Otherwise, the information is equal for both parts.

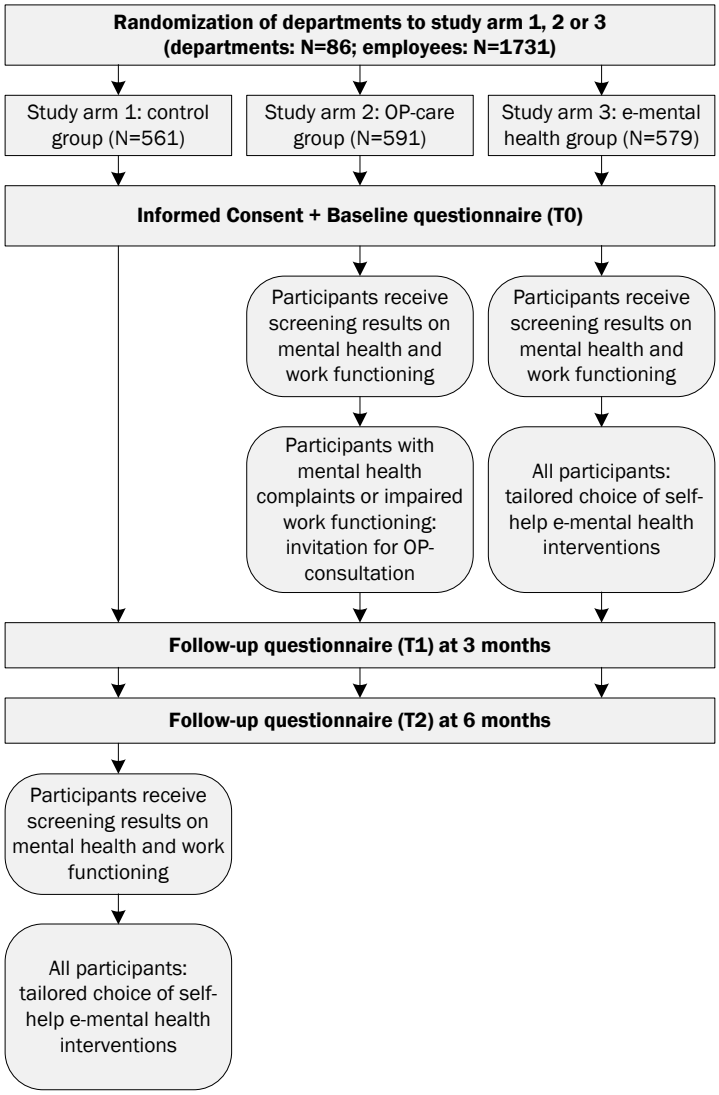


Figure 1. Study design.

Setting

The study will be performed in one Dutch academic medical center, a hospital with 7000 employees and 1102 beds. The organization has its own in-company occupational health service. Each of the different divisions in the medical center has one OP assigned for its occupational health care. In the Dutch occupational health care system, OPs fulfill preventive tasks, have the duty of detecting occupational diseases and provide (return-to-work)

counseling for sick-listed employees. In the Netherlands, if they are reported sick, workers are required to visit their OP for independent judgment of sick leave and for return-to-work guidance. Furthermore, all workers can make use of the free accessible consulting hour for employees with questions about work and health.[23] According to the in-company occupational health service, usage of the free accessible consultations by employees is limited.

Study population

The research population includes all nurses, including surgical nurses and anesthetic nurses, and allied health professionals working at one Dutch academic medical center. In total, 1731 nurses and allied health professionals work in the medical center in 86 different wards, including outpatient wards. Employees who are sick-listed at the start of the study and expected to be on sick leave for more than two weeks are excluded from the study, as they will in any case consult an OP.

Recruitment of participants

For a successful intervention study in a work setting, all layers of the organization must fully support the study. As we received endorsement from the board of directors, the nurse directors and the workers' counsel to perform the study, the likelihood for the departments to accept participation is high and all departments are expected to participate. During the enrollment period, information will be given on the hospital intranet, posters will be put up in the hospital building and flyers will be given out to promote the study.

Prior to the recruitment of individual employees, all wards will be informed about the Mental Vitality @ Work study by personal letters to the nurse directors and head nurses and the managers of allied health professionals. Subsequently, the individual workers will be informed by a letter to their home address with detailed information about the study procedure and about the safety and privacy of the individuals. Ten days later, an e-mail will be sent to the employees' work-email account, encompassing study information, a link to the online questionnaire and a personal log-in name and password. Agreeing with the informed consent, which is shown online prior to the questionnaire, is a prerequisite for starting the questionnaire. It will be possible for participants to log in on the website at any time and from any computer. It is also possible to log out at any point during the survey and to continue after logging in again. In the four weeks after the invitation for the study, three reminders will be sent to employees who have not yet started or completed the questionnaire.

Interventions

OP-care

The first WHS mental module strategy encompasses an online screening for mental health complaints and work functioning problems plus an optional consultation with an OP for employees with mental health problems and/or work functioning problems. Directly after

filling out the screening questionnaire, participants will automatically receive digital feedback on the screening results. Participants who score above a cut-off point for either the mental health complaints, work functioning impairments or both will receive an invitation from the in-company occupational health service for a face-to-face consultation with their OP scheduled within two weeks after filling out the screening questionnaire at baseline. The consultation is voluntary and supervisors of employees will not be informed about the invitation or the content of the consultation with the OP.

In order to structure the consultation of the OP, a seven-step protocol will be applied. The seven steps are: 1) discussing expectations; 2) discussing screening results and characteristics of mental health/work functioning complaints; 3) discussing possible causes in the private, work and medical situation and consequences for performing the work; 4) giving a diagnosis and offering a rationale; 5) giving advice for reduction of health complaints and for the improvement of work functioning and the prevention of incidents at work and discussing communication with the supervisor; 6) discussing a possible follow-up trajectory or referral to other care givers; and 7) summarizing the consultation.

This protocol closely follows the care as usual of the OPs. It was developed by means of interviews with the five participating OPs and based on the evidence-based guideline for OP's treatment of workers with mental health problems, which was developed by the Dutch Society of Occupational Medicine[11 24]. The main difference with the care as usual is the focus on identifying impairments in work functioning and giving advice on the improvement of work functioning and the prevention of consequences of impaired work functioning. All participating OPs were trained in using the protocol for the consultations.

E-mental health

The second WHS mental module strategy encompasses an online screening for mental health complaints and work functioning problems plus tailored advice on self-help e-mental health interventions. In this strategy, after filling out the screening questionnaire, feedback on results will be provided digitally. Workers with impaired work functioning will be digitally offered advice on how to improve their work functioning. Furthermore, an electronic health intervention trajectory will be offered to each participant to improve mental health and wellbeing. The trajectories offered for improvement of mental health will be tailored to the needs of the worker as assessed by the screening. The e-mental health interventions that can be offered are:

- Psyfit: aimed at the promotion of wellbeing. It is suitable for everyone, including healthy participants.[25] The effectiveness of Psyfit is currently being examined (Bolier, Bohlmeijer, Haverman, Boon, Kramer, Riper, unpublished).
- Strong at work (Sterk op je werk): aimed at gaining insight into work stress and at learning skills to cope with work stress.

- Colour your life (Kleur je leven): aimed at tackling depressive symptoms. Research has shown Colour your life to be (cost) effective.[26-31]
- Don't Panic Online (Geen Paniek Online): aimed at reducing panic symptoms for sub-clinical and mild cases of Panic Disorder. This intervention is based on Don't Panic (*Geen Paniek*), a face-to-face group course for sub-clinical and mild panic symptoms. Don't Panic has proven to be cost-effective.[32-35] The effectiveness of this online intervention is currently being studied.[36]
- Drinking less (Minder drinken): aimed at reducing risky alcohol drinking behavior and shown to be effective.[37 38]

The e-mental health interventions are self-help programs on the internet aimed at reducing specific mental health complaints or enhancing wellbeing. The programs are mainly based on cognitive behavioral therapy principles and combine a variety of aspects, for instance, advice, weekly assignments, the option of keeping a diary and a forum to get in contact with others who have similar complaints. The self-help e-mental health interventions were developed by the Trimbos-institute. E-mental health programs have been shown to be effective at improving impaired mental health[28 29 37-39] and at enhancing wellbeing.[40 41]

Control group

In the control arm, participants will fill out the baseline questionnaire; however, results of the screening-questionnaires will not be reported back to participants, and no further interventions will be advised at baseline. As compensation, participants in the control arm will receive their personal screening results together with a tailored choice for a self-help e-mental health intervention six months after baseline, which is identical to the intervention in the E-mental health arm at baseline.

For ethical reasons, a suicide-risk protocol is implemented in all study arms. Participants identified as being at high risk of suicide will receive immediate feedback on their self-reported suicide risk. They will be advised to seek help instantly, and they are asked to choose between either contacting their general practitioner or receiving an invitation for an urgent consultation with their OP.

Co-interventions

To our knowledge, no co-interventions on the organization or ward level aimed at mental health or work functioning improvement will be taking place in the medical center at the time of this study.

Measures

Screening instruments used at baseline

Impaired work functioning

Impaired work functioning will be measured using the job-specific Nurses Work Functioning Questionnaire (NWFQ) (Gärtner, Nieuwenhuijsen, van Dijk, Sluiter, unpublished). The NWFQ aims to measure impaired work functioning due to CMDs in nurses and allied health professionals. This 50-item self-report questionnaire consists of seven subscales: 1) *Cognitive aspects of task execution and general incidents*; 2) *Impaired decision making*; 3) *Causing incidents at work* (not applicable for allied health professionals); 4) *Avoidance behavior*; 5) *Conflicts and annoyances with colleagues*; 6) *Impaired contact with patients and their family*; and 7) *Lack of energy and motivation*. Cronbach's alphas vary between 0.70 and 0.94. Response formats vary between 5-category and 7-category scales; however, the number of categories is the same for all items of one subscale. The content of the response scales varies between Likert-type scales (0 = *totally disagree* to 6 = *totally agree*; 0 = *disagree* to 4 = *agree*; 0 = *no difficulty* to 6 = *great difficulty*), relative frequency categories (0 = *almost never* to 6 = *almost always*; *0* = *almost never* to 4 = *almost always*), and absolute frequency categories (0 = *not once* to 6 = *in general more than once a day*). Sum scores of the subscales range from 0-100. As yet, no validated cut-off scores exist for this fairly new questionnaire. Based on prior data of the study population (Gärtner, Nieuwenhuijsen, van Dijk, Sluiter, unpublished), cut-off values were calculated according to the following principle. Sumscores on the different subscales can lead to three categories: green, orange, and red. Therefore, two cut-off values are set, at the 67th percentile and at the 75th percentile of participants with mental health complaints. In two of the 7 subscales (subscale 2 and 4) the cut-off values for orange and red were identical due to little variation; in this case, cut-off values were set at the 75th and 85th percentiles of participants with mental health complaints. For the total NWFQ, a red score on one subscale or three or more orange scores will lead to case identification of impaired work functioning. In the prior dataset, this resulted in 31% of the total sample.

Distress

Distress will be measured with the distress subscale of the Four-Dimensional Symptoms Questionnaire (4DSQ).[42 43] The 16-item questionnaire uses a 5-point response scale (0 = *no*, 4 = *very often*) and has a Cronbach's alpha of 0.90.[43] For case identification, a cut-off point of ≥ 11 will be applied.[44]

Need for recovery

Early symptoms of work-related fatigue will be measured using the Need for recovery subscale of the Dutch Experience and Evaluation of Work (Dutch: VBBA) questionnaire.[45] The 11-item questionnaire with dichotomous response categories (*yes, no*) has a Cronbach's alpha of

0.86.[46-48] A cut-off point of ≥ 6 will applied. This gives a sensitivity of 0.72 and a specificity of 0.79.[49]

Alcohol use

To measure risky drinking behavior, the 3-item AUDIT-C will be used. The three items ask for frequency of specific drinking behavior, varying in formulations for the items.[50] Two items have a 5-point response scale, and 1 item has a 6-point response scale. The cut-off score is ≥ 5 for men with a sensitivity of 90.9 and specificity of 68.4 and ≥ 4 for women with a sensitivity of 92.4 and specificity of 74.3.[51]

Depression and Anxiety

Depression and anxiety will both be measured with the corresponding subscales of the Brief Symptom Inventory (BSI).[52] Each subscale has six items with a 5-point response scale (0 = *not at all*, 4 = *extremely*). Cronbach's alphas are 0.87 for both scales.[52] For both subscales, mean scores of ≥ 0.42 are used for case identification, with a sensitivity of 0.86 and a specificity of 0.66 for depression and a sensitivity of 0.83 and a specificity of 0.62 for anxiety.[53]

Suicide risk

One item of the BSI depression subscale asks for suicidal thoughts. An answer on this item in one of the upper two response categories (*rather a lot* or *extremely*) will identify a person as being at high risk for suicide.

Panic disorder

The panic module of the Patient Health Questionnaire (PHQ-15) will be used for the assessment of panic disorders; however, it will only be used in participants identified as having anxiety complaints.[54] The 15 items have dichotomous answering categories (*yes, no*) and a Cronbach's alpha of 0.80.[55] For case identification, we use the following procedure: two answers affirmative on the first four items plus four symptoms affirmative on the following 11 items. This identification procedure has a sensitivity of 0.91 and a specificity of 0.88.[56]

Post traumatic stress disorder

Post traumatic stress disorder is measured by the Schok Verwerkings Lijst (SVL),[57] a Dutch translation of the Impact of Event Scale.[58] The 15 items can be answered on a 4-point response scale (0 = *not at all*, 3 = *often*). Van der Ploeg et al.[59] found a Cronbach's alpha of 0.94 in a work-related sample. A cut-off point of ≥ 26 is applied.[60]

Work relatedness of mental health complaints

Work relatedness of mental health complaints is measured by one item: "Do you think that your work has negative consequences for your mental health?" This self-formulated item has a dichotomous response scale (*yes, no*).

At T2, the same screeners will be used in the control arm.

Primary outcomes measured at baseline, three month follow-up and six month follow-up

Substudy 1

The primary study parameter of the comparison between the OP-care arm and the control arm is help-seeking behavior. It regards formal help sources that the subject has used during the past three months. In the operationalization of formal help sources, 11 help sources are presented (i.e., psychologist, psychiatrist, general practitioner, OP, physiotherapist, supervisor, coach, in-company social worker, social worker, religious counselor, alternative therapeutic treatments). The list of help sources is developed in analogy with earlier studies on help-seeking behavior.[61-64]

The outcome measure help-seeking behavior is dichotomized into 'did seek formal help' for participants who had made use of any of the 11 caregivers and 'did not seek formal help' if none of the 11 caregivers were visited.

Substudy 2

The primary outcome measure of the comparison of the E-mental health arm with the OP-care and the control arm is work functioning, operationalized as job-specific impairments in work functioning. It will be measured using a total score of the Nurses Work Functioning Questionnaire (NWFQ).

Secondary outcomes measured at baseline, three month follow-up and six month follow-up

Secondary outcomes of both Substudy 1 and Substudy 2 are mental health complaints, and absenteeism. The secondary outcomes that are only measured for Substudy 1 are work functioning and additional help-seeking information (intention to seek help, work as content of the consultation, frequency of visits, and informal help-seeking behavior). The secondary outcomes that are only measured for Substudy 2 are work ability, turnover intention, wellbeing, and work productivity.

Mental health complaints

Mental health complaints are operationalized as the six mental health complaints screened for (i.e., distress, need for recovery, alcohol use, depression, anxiety and posttraumatic stress disorder). These are measured as described above.

Absenteeism

Three items from the Productivity and Disease Questionnaire (PRODISQ) Module C are used to measure absenteeism from work. Absenteeism is operationalized as number of days on sick leave during the last three months and number of periods of sick leave during the last three months.[65]

Work functioning

Work functioning will be measured by the NWFQ as described above.

Additional information on help-seeking behavior

Additional information concerning help-seeking behavior will be used as a secondary outcome measure, which includes 1) intention to seek help, assessed for the 11 formal help sources, 2) work as content of the consultation of various caregivers, 3) frequency of visits to the caregivers and 4) informal help-seeking behavior towards family or friends.

Work ability

Work ability will be assessed with the first item of the Work Ability Index (WAI).[66] This item concerns the evaluation of current work ability compared to their lifetime best on an 11 point scale (0 = *completely unable to work*, 10 = w*ork ability at its best*).

Turnover intention

Turnover intention will be assessed by one item on plans to seek for a job outside of the present organization during the next year. The item can be answered on a dichotomous response scale (*yes, no*).

Wellbeing

Wellbeing is measured with three questionnaires measuring different concepts.

The Mental Health Continuum-Short Form (MHC-SF) is a 14-item self-report questionnaire on wellbeing in the categories 'languishing', 'moderate' and 'flourishing'.[67] The MHC-SF measures hedonistic wellbeing as well as psychological and social wellbeing. Participants rate the items on a 6-point scale (0 = *never*, 5 = *every day*). The MHC-SF has shown good internal consistency (> 0.80) and discriminant validity.[68 69]

The WHO-5 wellbeing scale contains five positively formulated items on mental health. Participants are asked to rate the items using a 6-point scale (0 = *never*, 5 = *all of the time*). The WHO-5 has been validated in different populations with an acceptable internal consistency (Cronbach's alpha 0.84).[70]

The Utrecht Work Engagement Scale (UWES-9, short-form) measures engagement at the workplace. It is a 9-item scale, and items are scored on a 7-point rating scale (0 = *never*, 6 =

always). Cronbach's alpha of the UWES-9 varied between 0.85 and 0.92 across 10 different countries, including the Netherlands.[71]

Work productivity

Three items from the Productivity and Disease Questionnaire (PRODISQ) Module E are used to measure productivity losses due to presenteeism. The three items refer to the last work day, and they assess the amount of inefficient job performance, the quality loss of the work, and, if applicable, the reason for productivity loss.[65]

Independent measures at baseline

As independent measures, we assess demographic characteristics, job characteristics and psychosocial work characteristics at baseline. Demographic characteristics, gender, age (in years), civil status (five categories), and ethnic background (three categories) will be assessed with self-formulated questions. As job characteristics, we will measure the occupation, nursing specialty (if applicable), work experience in years, work hours per week, and type of labor contract. Psychosocial work characteristics will include job demands, job control, social support at work from the supervisor, and social support at work from colleagues, which will each be measured by one self-formulated item on a visual analogue scale (VAS) (0 = *not*, 100 = *to great extent*). Additionally, one item will be added for conflicts at work with the supervisor or with colleagues. As possible prognostic factors for help-seeking behavior, we include gender,[72-74] civil status,[75] informal help-seeking towards family or friends,[76] and previous experiences with mental health care, which is operationalized as having friends or family who were treated by a psychologist/psychiatrist at any time, or having been treated by a psychologist/psychiatrist himself/herself at any time.[77]

Process indicators measured at three month follow-up

Process indicators for the feasibility evaluation of the WHS will be measured at three month follow-up (T1) and include 1) participants' compliance in both the OP-care and E-mental health arm; 2) participants' satisfaction; 3) adherence of OP to the protocol; and 4) satisfaction of OP. Participants' compliance will be assessed by objective data on response rate to the study, percentages of participants who made use of the invitation for an OP consultation or the e-mental health interventions (by track and trace); moreover, based on self-report data, the percentages of participants who followed the advice given by the OP or during the e-mental health intervention. Satisfaction of participants will be measured by self-report data on satisfaction with the provided feedback, satisfaction with the invitation for the OP consultation or the e-mental health intervention, satisfaction with the consultation by the OP or e-mental health intervention itself, and satisfaction with the advice given by the OP or given in the e-mental health intervention, including their perceived effectiveness. In the OP-care arm, protocol adherence of the OPs will be assessed by means of a checklist for each protocol step,

which the OP has to fill out after each consultation with a participant of the WHS mental module. The OP's satisfaction and experiences with the WHS mental module will be assessed in a group interview after the three month follow-up.

Economic evaluation indicators at baseline, three month follow-up and six month follow-up

The cost-effectiveness of the WHS mental module will be assessed from a societal perspective. Differences in effect – job-specific impairments in work functioning – will be compared with differences in costs – costs due to intervention and health care and costs stemming from productivity losses in paid work.

The employer's perspective will be considered in a cost-benefit analysis by comparing the costs of occupational health care (including the WHS mental module) with the costs due to productivity losses in paid work. Health care utilization will be measured by the Trimbos/iMTA Cost Questionnaire for Psychiatric Illness (TiC-P).[61] Questions on occupational healthcare utilization will be added to this questionnaire for the purpose of this study. The Productivity and Disease Questionnaire (PRODISQ) will be used to measure productivity losses due to absenteeism and presenteeism (inefficient job performance).[65]

Sample Size

Substudy 1

In a study by Isaaksson Ro[78] on help-seeking behavior in nurses with burnout, the formal help-seeking increased from 17% to 34%. Differences between the percentages of participants having sought formal help between the two study arms will be examined using a Chi-square test. For an increase of 17% with alpha = 0.05 (2-tailed) and a power of (1-beta) = 0.80, power calculation using the Nquery Advisor software results in 114 participants with mental health complaints for each of the two arms. Based on data from a prior study in this population (Gärtner, Nieuwenhuijsen, van Dijk, Sluiter, unpublished), we assume that 50% of the population will have impairments in either mental health, work functioning or both. Thus, for a comparison of workers screened positive in the control arm and the OP-care arm, 228 participants in each arm are necessary. Randomization will take place at the ward level; however, we do not expect any correlation between wards in the help-seeking behavior of their employees. Therefore, no inflation factor is used in the power calculation for Substudy 1 with the outcome measure help-seeking behavior. With an expected loss-to-follow-up of 10%, we must start the trial with N = 228/0.90 = 254 per condition at baseline.

Substudy 2

The trial is powered to detect a clinically significant effect, defined as at least 0.33 standard units when the (primary) outcome is transformed into a standardized effect size, also known as Cohen's d or the standardized mean difference. Lipsey and Wilson[79] conducted a second-order meta-analysis of psychological, educational and behavioral interventions and found that

for these interventions, d = 0.33 to be corresponding with the lower bound of a medium effect size. We will conduct tests at alpha = 0.05 (2-tailed) and a power of (1-beta) = 0.80. Using Stata, it is shown that N = 145 per condition is required. For the primary outcome measure of this substudy, work functioning, no information on probable difference on ward-level exists. But as a precaution, we compensate for possible cluster effects introduced into the data because of randomization on ward-level. For cluster correction, we must multiply by a factor 1.3, which returns 145*1.3 = 189 per condition. Assuming a loss-to-follow-up of 10%, we must start the trial with N = 189/0.90 = 210 per condition at baseline.

In sum, the required minimum number of participants is 254 for the control arm, 254 for the intervention arm 2 (according to calculations for part 1) and 210 for the intervention arm 3 (according to calculations for Substudy 2); thus, in total, 718 participants are required for all three arms. We expect a response rate of about 45%; thus, 1596 employees must be invited to recruit the required 718 participants. As we will include 1731 employees, the source population is large enough to meet the needed sample size.

Randomization and blinding

In this controlled trial, cluster randomization will be performed at the ward level. The argumentation for cluster randomization is two-fold. First, it prevents contamination effects between participants working in the same department. Second, it is in accordance with the practice of WHS, which is usually conducted for all workers in a department. The randomization procedure will take place before the inclusion of the individual participants. In the randomization, we will stratify for ward size. Randomization will be performed using block randomization with three departments in each block. To guarantee concealment of allocation, the randomization will be performed by one researcher (KN) who is not involved in the practical recruitment of participating employees, using the computer software program Nquery Advisor.

Researchers, managers of participating departments and OPs will not be blinded for the group allocation. However, as we have a prerandomization procedure with incomplete-double-consent design without mentioning the use of a reference group in the experimental groups and vice versa,[80] the head (nurses) of wards and the individual employees will receive only information that is applicable to the study-arm of their wards.

Statistical analyses

The baseline data and data of the primary and secondary parameters will be presented using descriptive statistics. The effectiveness of the intervention on the primary and secondary outcome measures will be analyzed on the employee level following the intention-to-treat-principle.

Effect evaluation

To study the effect on dichotomous outcome measures Chi-square tests will be used; thus, to test differences in proportions of subjects who score positive on the outcome measure between the study arms for each time of measurement. Change of proportion of employees in outcomes at follow-up (T1 and T2) will be analyzed using Generalized Estimated Equations (GEE), with wards and participants as cluster variables (where appropriate) and study-arm, time and their interaction (study-arm x time) as co-variates under the assumptions of an exchangeable working correlation matrix. Effects of demographic characteristics and prognostic factors on dichotomous outcomes measures will also be analyzed using Generalized Estimated Equations (GEE).

Effects of continuous outcome measures will be analyzed using multiple regression analysis. A multilevel analysis of variance will be conducted (GLM mixed models, repeated measurements), with ward as the primary hierarchical level and participants as the secondary hierarchical level (where appropriate). Effects of demographic characteristics and prognostic factors on continuous outcome measures will also be analyzed using multiple regression analysis.

Cost-effectiveness evaluation

For the cost-effectiveness evaluation, the incremental cost-effectiveness ratio (ICER) will be calculated by comparing the differences in costs of health care utilization and productivity losses for each WHS strategy with the difference in effect on job-specific work impairments of both strategies. The index year for health care costs will be 2011. Productivity losses will be assessed using the human capital approach. Analyses will include cost-effectiveness planes and acceptability curves. Ancillary analyses (i.e., incremental net-benefit regression analysis) will identify subgroups of workers (e.g., participants with and without impaired work functioning or mental health) who derive particular benefit from the intervention.

In the cost-benefit analysis, the employer's perspective will be considered by comparing the costs of offering the WHS modules with the costs of productivity losses due to sickness absence (absenteeism) and working less efficiently while at work (presenteeism) for both WHS strategies separately.

Process evaluation

Participant compliance and participant satisfaction as well as adherence of OP will be presented in proportions. Satisfaction of OPs will be assessed in terms of strengths and suggestions for improvement.

Ethical considerations

There are no risks associated with participating in the Mental Vitality @ Work study. Confidentiality is guaranteed during the whole study for the employees of all study arms, as no

information about the screening or the interventions will be provided to others, such as supervisors. Furthermore, the study participants of all study-arms retain unrestricted access to care as usual if requested. Employees and their supervisors are still free to call in any occupational health care in the medical center if they wish to do so.

DISCUSSION

The health care service is a sector with special risks for the development of mental health complaints. In turn, in this sector, impaired mental health can have serious consequences for the workers and their patients. A WHS mental module might be an effective preventive action to promote and monitor good (mental) health and work performance in the aging workforce. The aim of the Mental Vitality @ Work study is to test the effectiveness of two strategies for a WHS mental module for nurses and allied health professionals. This paper describes the protocol for a three-arm RCT in which the effectiveness of the two strategies for a WHS mental module will be evaluated. First, the effect on help-seeking behavior for the OP-care arm compared to a control arm and, second, a comparison of the effect on work functioning for the E-mental health arm with the OP-care arm and the control arm. Additionally, an economic evaluation of both procedures will be evaluated from both a societal and employer perspective.

WHS is a well-developed strategic concept to protect workers against health risks and to monitor and enhance their work functioning. Mental modules for WHS have been developed in some sectors, e.g., the police sector.[17] In these studies, the identification of workers in need of health care intervention was solely based on the mental health status. The innovative aspect of our approach is that, in addition to screening for mental health problems, a screening for work functioning problems is carried out. The identification of work functioning problems in workers with mental health complaints yields input for the kind of intervention needed to enhance work functioning and to prevent more serious consequences such as incidents at work. In line with this, our approach differs from other mental health screenings in the work setting, because we will test the effectiveness of the WHS mental module both at enhancing work functioning and at improving mental health.

Another innovative aspect of our study is the included e-mental health interventions. Although the effects of e-mental health interventions on mental health outcomes appear promising, applying them in the context of WHS in a specific working population is a new approach. An advantage of this context is that the e-mental health intervention can be tailored to the mental health outcome of the screening that precedes the offered interventions.

Methodological considerations

One strength of our RCT-design is the cluster randomization with pre-randomization. Applying a WHS procedure to a ward as a whole is not only in line with WHS in common occupational health service practice, but it also reduces contamination of employees. The pre-randomization approach allows blinding of participants for information of the other study arms. Still, contamination effects due to communication and occasional switching between wards cannot be ruled out completely, as the study is conducted in one organization.

One methodological issue to be considered regards our choice for not applying an inflation factor for cluster correction in Substudy 1. This choice is based on two arguments. First, we do not expect any systematic differences between the hospital wards in differences on health seeking behavior of their individual workers, which makes cluster correction illogical. Furthermore, we do not expect any noteworthy differences between the study arms in baseline characteristics, due to the large amount of clusters (N = 86). For work functioning, the primary outcome measure of Substudy 2, the possibility of systematic differences between the wards, is more likely. It is conceivable that improvement in work functioning, e.g., decision making, is more difficult for workers of one ward than workers of another ward, due to differences in work context. Therefore, a cluster correction is applied on Substudy 2.

We expect the external validity of this study to be high, as the study is encompassed in a real-life setting. Furthermore, in the set-up of the interventions, good feasibility is allowed for by using input of (nurse) managers and the occupational health service that provides the OP-care. The protocol for the OP consultations is developed based on interviews with the OPs and follows care as usual closely.

Impact of results

The output of the Mental Vitality @ Work study will be two-fold. First, two WHS mental modules for nurses and allied health professionals will be delivered. Based on results on the effectiveness together with results on the process evaluation, a WHS mental module for nurses and allied health professionals could be implemented and used on a regular basis by occupational health services in academic medical centers. A WHS mental module can be used as a stand-alone intervention or as part of a broader WHS program. With minor modifications, the module can be adapted to the context of other healthcare organizations.

Secondly, the proposed study will yield valuable knowledge on the effectiveness and cost-effectiveness of a WHS mental module. If it is effective in terms of costs and improvement of adequate help-seeking, work functioning, and improved mental health, the procedure for a WHS mental module will possibly be used as a blue-print and contribute to the development of WHS mental modules in other sectors. It also might promote the use of WHS in the Netherlands. Results of the study will become available in 2012.

REFERENCES

1 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
2 Hilton MF, Scuffham PA, Sheridan J et al. Mental ill-health and the differential effect of employee type on absenteeism and presenteeism. *J Occup Environ Med* 2008;50:1228-43.
3 Lerner D, Adler DA, Chang H et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46:S46-S55.
4 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
5 McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev* 2009;29:243-59.
6 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.
7 Gershon RR, Stone PW, Zeltser M et al. Organizational climate and nurse health outcomes in the United States: a systematic review. *Ind Health* 2007;45:622-36.
8 Sundin L, Hochwalder J, Bildt C et al. The relationship between different work-related sources of social support and burnout among registered and assistant nurses in Sweden: a questionnaire survey. *Int J Nurs Stud* 2007;44:758-69.
9 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
10 Nieuwenhuijsen K, Verbeek JH, de Boer AG et al. Validation of performance indicators for rehabilitation of workers with mental health problems. *Med Care* 2005;43:1034-42.
11 van der Klink JJ, Blonk RW, Schene AH et al. Reducing long term sickness absence by an activating intervention in adjustment disorders: a cluster randomised controlled design. *Occup Environ Med* 2003;60:429-37.
12 Coles ME, Coleman SL. Barriers to treatment seeking for anxiety disorders: initial data on the role of mental health literacy. *Depress Anxiety* 2010;27:63-71.
13 Dahlberg KM, Waern M, Runeson B. Mental health literacy and attitudes in a Swedish community sample - investigating the role of personal experience of mental health care. *BMC Public Health* 2008;8:8.
14 International Labour Organization (ILO). Technical and ethical guidelines for workers' health surveillance. Geneva: ILO, 1998.
15 Koh D, Aw TC. Surveillance in occupational health. *Occup Environ Med* 2003;60:705-10, 633.
16 Plat MJ, Frings-Dresen MH, Sluiter JK. A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *Int Arch Occup Environ Health* 2011;84:839-57.
17 Winwood PC, Tuckey MR, Peters R et al. Identification and measurement of work-related psychological injury: piloting the psychological injury risk indicator among frontline police. *J Occup Environ Med* 2009;51:1057-65.
18 Levin SM, Herbert R, Moline JM et al. Physical health status of world trade center rescue, recovery workers, volunteers-New York City, July 2002-August. *Morb Mortal Week Report* 2004;53:807-12.
19 Wright KM, Huffman AH, Adler AB et al. Psychological screening program overview. *Mil Med* 2002;167:853-61.
20 Netherlands Society of Occupational Medicine (NVAB). Guidance Workers' Health Surveillance [In Dutch: Leidraad Preventief Medisch Onderzoek]. Utrecht, the Netherlands: NVAB, 2005.
21 Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004;328:702-8.
22 Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010;63:834-40.
23 Hulshof CT, Frings-Dresen MH. International OH systems Part 2: occupational health deliveray in the Netherlands. *Occup Health Work* 2010;6:19-23.
24 van der Klink JJ. Guideline: management of mental health problems of workers by occupational physicians [In Dutch: Richtlijn: Handelen van de bedrijfsarts bij werkenden met psychisch problemen]. Utrecht, the Netherlands: NVAB [Netherlands Society of Occupational Medicine], 2007.
25 Haverman M, Bohlmeijer E, Bolier L. Psyfit, an online intervention to promote mental health [In Dutch: Psyfit, een internet interventie ter bevordering van de mentale gezondheid]. Utrecht: Trimbos-instituut, 2010.
26 de Graaf LE, Gerhards SA, Arntz A et al. Clinical effectiveness of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry* 2009;195:73-80.

27 de Graaf LE, Huibers MJ, Riper H et al. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. *J Affect Disord* 2009;116:227-31.
28 Spek V, Nyklicek I, Smits N et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. *Psychol Med* 2007;37:1797-806.
29 Spek V, Cuijpers P, Nyklicek I et al. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007;37:319-28.
30 Spek V, Cuijpers P, Nyklicek I et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. *Psychol Med* 2008;38:635-9.
31 Warmerdam L, van SA, Twisk J et al. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. *J Med Internet Res* 2008;10:e44.
32 Meulenbeek P, Willemse G, Smit F et al. Early intervention in panic: randomized controlled trial and cost-effectiveness analysis. *Trials* 2008;9:67.
33 Meulenbeek P, Willemse G, Smit F et al. Effects and feasibility of a preventive intervention in sub-threshold and mild panic disorder: Results of a pilot study. *BMC Res Notes* 2009;2:4.
34 Meulenbeek P, Willemse G, Smit F et al. Early intervention in panic: pragmatic randomised controlled trial. *Br J Psychiatry* 2010;196:326-31.
35 Smit F, Willemse G, Meulenbeek P et al. Preventing panic disorder: cost-effectiveness analysis alongside a pragmatic randomised trial. *Cost Eff Resour Alloc* 2009;7:8.
36 van Ballegooijen W., Riper H, van SA et al. The effects of an Internet based self-help course for reducing panic symptoms - Don't Panic Online: study protocol for a randomised controlled trial. *Trials* 2011;12:75.
37 Kramer J, Riper H, Lemmers L et al. Television-supported self-help for problem drinkers: a randomized pragmatic trial. *Addict Behav* 2009;34:451-7.
38 Riper H, Kramer J, Smit F et al. Web-based self-help for problem drinkers: a pragmatic randomized trial. *Addiction* 2008;103:218-27.
39 Riper H, Smit F, van der Zanden R et al. E-mental health. High tech, high touch, high trust. Utrecht: Trimbos-instituut, 2007.
40 Mitchell J, Stanimirovic R, Klein B et al. A randomised controlled trial of a self-guided internet intervention promoting well-being. *Comput Hum Behav* 2009;25:749-60.
41 Seligman MEP, Steen TA, Park N et al. Positive Psychology Progress: Empirical Validation of Interventions. *Am Psychol* 2005;60:410-21.
42 Terluin B. The Four Dimensional Symptom Questionnaire (4DSQ) [In Dutch: De Vierdimensionele Klachtenlijst (4DKL) in de huisartspraktijk]. *De Psycholoog* 1998;33:18-24.
43 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.
44 van Rhenen W., van Dijk FJ, Schaufeli WB et al. Distress or no distress, that's the question: A cutoff point for distress in a working population. *J Occup Med Toxicol* 2008;3:3.
45 van Veldhoven M, Meijman TF. The measurement of psychosocial job demands with a questionnaire: the questionnaire on the experience and evaluation of work (QEEW) [In Dutch: Het meten van een psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van arbeid (VBBA)]. Amsterdam: Dutch Institute for Working Conditions, 1994.
46 de Croon EM, Sluiter JK, Frings-Dresen MH. Psychometric properties of the Need for Recovery after work scale: test-retest reliability and sensitivity to detect change. *Occup Environ Med* 2006;63:202-6.
47 Sluiter JK, de Croon EM, Meijman TF et al. Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occup Environ Med* 2003;60 Suppl 1:i62-i70.
48 van Veldhoven MJ, Sluiter JK. Work-related recovery opportunities: testing scale properties and validity in relation to health. *Int Arch Occup Environ Health* 2009;82:1065-75.
49 Broersen JPJ, Fortuin RJ, Dijkstra L et al. Monitor Arboconvenanten: kengetallen en grenswaarden [Monitor occupational health and safety: key indicators and limits]. *TBV* 2004;12:100-4.
50 Bush K, Kivlahan DR, McDonell MB et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998;158:1789-95.
51 Gual A, Segura L, Contel M et al. Audit-3 and audit-4: effectiveness of two short forms of the alcohol use disorders identification test. *Alcohol Alcohol* 2002;37:591-6.
52 de Beurs E. Brief Symptom Inventory (BSI): Manual [In Dutch: Brief Symptom Inventory (BSI): Handleiding]. Leiden: Pits Publishers, 2006.

53 de Beurs E, Zitman FG. The Brief Symptom Inventory (BSI): The reliability and validity of a brief alternative of the SCL-90 [In Dutch: De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90]. *Maandblad GeestelijkeVolksgezondheid* 2006;61:120-41.
54 Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999;282:1737-44.
55 Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;64:258-66.
56 Lowe B, Grafe K, Zipfel S et al. Detecting panic disorder in medical and psychosomatic outpatients: comparative validation of the Hospital Anxiety and Depression Scale, the Patient Health Questionnaire, a screening question, and physicians' diagnosis. *J Psychosom Res* 2003;55:515-9.
57 Brom D, Kleber RJ. The Impact of Event Scale [In Dutch: De Schok Verwerkings Lijst]. *Nederlands Tijdschrift voor de Psychologie* 2011;40:164-8.
58 Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosom Med* 1979;41:209-18.
59 van der Ploeg E, Mooren TT, Kleber RJ et al. Construct validation of the Dutch version of the impact of event scale. *Psychol Assess* 2004;16:16-26.
60 Chemtob CM, Tomas S, Law W et al. Postdisaster psychosocial intervention: a field study of the impact of debriefing on psychological distress. *Am J Psychiatry* 1997;154:415-7.
61 Hakkaart-van Roijen L, van Straten A, Donker M. Trimbos/iMTA questionnaire for costs associated with psychiatric illness (TIC-P). Rotterdam: iMTA, 2002.
62 Rickwood D, Deane FP, Wilson CJ et al. Young people's help-seeking for mental health problems. *AeJAMH* 2005;4:1-34.
63 Walters K, Buszewicz M, Weich S et al. Help-seeking preferences for psychological distress in primary care: effect of current mental state. *Br J Gen Pract* 2008;58:694-8.
64 Wilson CJ, Deane FP, Ciarrochi J et al. Measuring Help-Seeking Intentions: Properties of the General Help-Seeking Questionnaire. *Can J Counsell* 2005;39:15-28.
65 Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5:23-8.
66 Tuomi K, Ilmarinen J, Jahkola A et al. Work Ability Index. 2nd revised ed. Helsinki: Finnish Institute of Occupational Health, 1998.
67 Keyes CL. The mental health continuum: from languishing to flourishing in life. *J Health Soc Behav* 2002;43:207-22.
68 Lamers SMA, Westerhof GJ, Bohlmeijer ET et al. Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF). *J Clin Psychol* 2011;67:99-110.
69 Westerhof GJ, Keyes CLM. Mental health is more then the absence of disease [In Dutch: Geestelijke gezondheid is meer dan de afwezigheid van ziekte]. *Maandblad Geestelijke Gezondheid* 2008;10:808-20.
70 Bech P, Olsen LR, Kjoller M et al. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale. *Int J Methods Psychiatr Res* 2003;12:85-91.
71 Schaufeli WB, Bakker AB, Salanova M. The Measurement of Work Engagement With a Short Questionnaire: A Cross-National Study. *Educ Psychol Meas* 2006;701-16.
72 Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: literature review. *J Adv Nurs* 2005;49:616-23.
73 Judd F, Komiti A, Jackson H. How does being female assist help-seeking for mental health problems? *Aust N Z J Psychiatry* 2008;42:24-9.
74 Moller-Leimkuhler AM. Barriers to help-seeking by men: a review of sociocultural and clinical literature with particular reference to depression. *J Affect Disord* 2002;71:1-9.
75 Gudmundsdottir G, Vilhjalmsson R. Group differences in outpatient help-seeking for psychological distress: results from a national prospective study of Icelanders. *Scand J Public Health* 2010;38:160-7.
76 Rudell K, Bhui K, Priebe S. Do 'alternative' help-seeking strategies affect primary care service use? A survey of help-seeking for mental distress. *BMC Public Health* 2008;8:207.
77 Schomerus G, Matschinger H, Angermeyer MC. The stigma of psychiatric treatment and help-seeking intentions for depression. *Eur Arch Psychiatry Clin Neurosci* 2009;259:298-306.
78 Isaksson Ro KE, Gude T, Tyssen R et al. A self-referral preventive intervention for burnout among Norwegian nurses: one-year follow-up study. *Patient Educ Couns* 2010;78:191-7.
79 Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am Psychol* 1993;48:1181-209.
80 Schellings R, Kessels AG, ter RG et al. Indications and requirements for the use of prerandomization. *J Clin Epidemiol* 2009;62:393-9.

7

The Mental Vitality @ Work study – Effects of a workers' health surveillance mental module for nurses and allied health professionals: a cluster-randomized controlled trial

Fania R. Gärtner
Karen Nieuwenhuijsen
Sarah M. Ketelaar
Frank J.H. van Dijk
Judith K. Sluiter

Submitted for publication.

ABSTRACT

Objectives: The effectiveness of a workers' health surveillance (WHS) mental module for nurses and allied health professionals was studied. It is expected to increase help-seeking behavior and improve work functioning and mental health.

Methods: A cluster randomized-controlled trial was performed. In total, 1,152 nurses and allied health professionals from 57 wards of one Dutch academic medical center were invited to participate. The WHS mental module included screening for impairments in work functioning and mental health with immediate feedback on the screening results online. Workers who were screened as positive were invited to visit their occupational physicians. The control group received no screening results or intervention. Outcomes were assessed at baseline, three months, and six months of follow-up. Effects with respect to the primary outcome, help-seeking behavior, were analyzed using generalized linear mixed models. Effects on the secondary outcomes, work functioning and mental health, were analyzed using linear mixed models.

Results: In the intervention group, 79% of participants were screened as positive for mental health complaints and/or work functioning impairments compared to 86% in the control group. A statistically significant interaction effect of study-group*time on help-seeking behavior was found ($p = 0.02$). In the intervention group help-seeking remained stable between baseline and three months follow-up, while it decreased for the control group. Towards six months follow-up, help-seeking behavior decreased more steeply in the intervention group than it did in the control group, resulting in even slightly less help-seeking behavior. Workers who received the WHS mental module showed less work functioning impairments over time compared with the control group ($p = 0.04$). No statistically significant effects on mental health complaints were found, except for an interaction effect of study-group*time on risky drinking behavior ($p < 0.01$).

Conclusion: The WHS mental module can be used effectively to stimulate help seeking at the occupational physician and to improve work functioning in workers with mild to moderate work functioning and/or mental health complaints.

Trial number: NTR2786

INTRODUCTION

Common mental disorders (CMD) have a large impact on work participation.[1] This group of disorders, which encompasses distress, adjustment, depressive, and anxiety disorders, is associated with a substantial effect on sickness-related absence and impairments in work functioning.[2 3] In healthcare, the adverse work effects of CMDs can be serious. For nurses, CMDs are related to higher risks for workers' and patients' safety, as the functioning at work of workers who stay at work with mental health complaints might be impaired.[4] Moreover, CMDs are highly prevalent in nurses.

When CMDs go untreated, they can result in the aggravation of the mental health complaints. The later workers receive help, the more difficult successful treatment becomes.[5] Although effective care for mental health complaints is available, mental health care is often not used by workers with CMDs.[6] One reason for no or late help seeking might be the lack of recognition of one's own mental health complaints.[7] In a study across six European countries, only 33% of adults with a mental health disorder perceived the need for mental health care.[8] Aside from the lack of recognition, negative attitudes toward health care are barriers to active help seeking.[9] Healthcare workers in particular seem to experience these barriers.[10] This group is used to provide care instead of receiving care and therefore finds it difficult to seek help for themselves.[11] Additionally, fear of stigmatization might play a role in the avoidance of active help seeking, especially in a work context.[12]

Preventive actions might be useful to identify workers with mild symptoms of CMDs and/or early signals of impaired work functioning to encourage active help seeking. One occupational health strategy for preventive actions is workers' health surveillance (WHS).[13 14] In WHS, the periodic monitoring of employees' health and work functioning aims to detect the negative health effects of work early, to enable timely interventions.[14] Internationally, WHS most often focuses on health effects. In the Netherlands, following the professional guidelines,[15] WHS focuses on the improvement of both health and work functioning. It consists of a screening to detect employees with (early) health or work functioning problems and interventions to improve complaints and to reduce the risks of aggravating any health issues. Little is known about WHS targeting mental health effects. In a recent literature review by Plat and colleauges,[16] only three of the 24 studies on WHS included aspects of mental health. WHS for factors related to mental health and its consequences for the work of nurses or other healthcare service professionals have not yet been evaluated scientifically.

The aim of the present study is to test a WHS mental module for nurses and allied health professionals. The WHS mental module includes an online screening and an invitation for a face-to-face preventive consultation with an occupational physician for participants who are screened as positive for impaired work functioning and/or mental health complaints. We expect that use of a screening in combination with personal feedback, as well as the invitation to consult with an occupational physician, stimulates active help-seeking behavior.

Furthermore, we expect that work functioning and mental health problems will improve in workers who receive the WHS mental module.

METHODS

Study design

This study is part of a cluster-randomized controlled trial on the effectiveness of two strategies for the administration of a WHS mental module for nurses and allied health professionals, the Mental Vitality @ Work study. The study design has been published.[17] The complete trial included two intervention groups and one control group. The present study focuses on the effect evaluation of one strategy with occupational physicians' care (WHS OP-care group). Researchers and managers of participating departments and occupational physicians were not blinded for the group allocation. However, as a pre-randomization procedure with an incomplete-double-consent design was applied, the head (nurses) of wards and the individual workers received only information that was applicable to the study groups on their wards.[18] Randomization was performed at the ward level. Randomization sequences were generated using the computer software program Nquery Advisor with a block size of three. To guarantee allocation concealment, randomization was performed by one researcher (KN) who was not involved in the practical recruitment of employees.

The Medical Ethics Committee of the Academic Medical Center Amsterdam declared that the study design did not require comprehensive ethical review. Individual informed consent was required online prior to the start of the baseline questionnaire.

Participants

The study population of the complete trial consisted of 1,731 nurses and allied health professionals employed at one Dutch academic medical center. Participants were included if they were not and not expected to be on sick leave (> 2 wks) at the start of the study. All eligible employees were invited to take part in the study. Randomization at the ward level (N = 86) resulted in a total of 561 employees being assigned to the control group and 591 employees being assigned to the WHS OP-care group, which resulted in the involvement of 57 wards.

Based on sample size calculation, the minimum number of participants was 508, 254 for each of the two study groups.[17] This calculation was based on the need for 114 participants per study group who would screen positive for work functioning impairments and/or mental health complaints and fill out all three questionnaires.

Procedure

In March 2011, the workers were invited by e-mail to fill out the electronic baseline questionnaire. It was possible to discontinue the questionnaire and complete it after logging in again. Three reminders were sent. Three and six months after baseline, those who had completed the baseline questionnaire were invited to complete the follow-up questionnaires.

The WHS mental module

The WHS mental module consisted of two parts: first, an online screening and personal feedback on the screening results; second, an invitation to consult with an occupational physician when screened as positive.

The screening consisted of an assessment of seven aspects of work functioning and six types of mental health complaints: distress, need for recovery, risky drinking behavior, depression and suicide risk, anxiety and panic, and post-traumatic stress disorder. The measurement instruments used for the screening are described in detail by Gärtner and colleagues.[17] After completing the screening questionnaire, the participants received immediate feedback on their results online.

Participants in the WHS OP-care group who were screened as positive received an invitation by letter to their home address for a face-to-face preventive consultation with their own occupational physician. The consultation was voluntary, and workers could reschedule or cancel it if they wished. Supervisors were not informed.

A consultation protocol was developed based on interviews with participating occupational physicians to elicit their current practice and on the Dutch evidence-based guidelines for occupational physicians' counseling of workers with mental health problems.[19] The seven-step protocol offered a standardized procedure for the preventive consultation. The protocol closely followed occupational physicians' care as usual for open consultations, which, in the Netherlands, are consultations initiated by the employee in contrast to the compulsory consultation in the context of absenteeism. The occupational physicians were required to pay special attention to the participants' recognition of their screening results. Occupational physicians received 3 h of training from the researchers on the use of the protocol.

Primary outcome measures

Help-seeking behavior

Help-seeking behavior regarded formal help sources that the subject had used during the past three months. Eleven help sources were presented: psychologist, psychiatrist, general practitioner, occupational physician, physiotherapist, supervisor, coach, in-company social worker, social worker, religious counselor, and alternative therapeutic treatments. This list was developed similarly to earlier studies on help-seeking behavior.[20 21] Objective data from the in-house occupational health service on worker visits for open consultations were used to correct the self-reported data on visits to the occupational physician.

Help-seeking behavior was dichotomized into 'did seek formal help' for participants who had made use of any of the 11 caregivers and 'did not seek formal help' if none of the 11 caregivers were visited.

Covariates

Four variables were considered as potential prognostic variables for the effect on help-seeking behavior and assessed at baseline. First, gender;[22] second, civil status (divorced at baseline or not);[23] third, informal help seeking toward family or friends (having sought help from family or friends in the past three months or not);[24] and fourth, previous experiences with mental health care (having friends or family who were treated by a psychologist/psychiatrist at any time or having been treated by a psychologist/psychiatrist himself/herself at any time).[25]

Secondary outcome measures

In deviation from the study protocol, the frequency of visits to caregivers and absenteeism were not studied as secondary outcome measures. These outcomes will be included in the economical analysis of the WHS mental module.

Additional information on help-seeking behavior

Additional information concerning help-seeking behavior included the following: 1) the intention to seek help (*yes/no*), as assessed for the 11 formal help sources; 2) informal help-seeking behavior toward family or friends (*yes/no*); and 3) discussion of work during consultation with various caregivers, as assessed for 9 of the 11 caregivers (not for supervisor or occupational physician) on a 5-point scale (*0 = no, 4 = the bulk of the consultation*). This score was dichotomized as follows: 0 and 1 were considered as 0 (*work was not a focus of the help sought*), and 2, 3, and 4 were considered as 1 (*yes, work was a focus of the help sought*).

Impaired work functioning

Impaired work functioning was measured using the job-specific Nurses Work Functioning Questionnaire (NWFQ).[26 27] In contrast to the description prior to the start of the trial, only six of the seven original NWFQ subscales were used as outcomes measure, as the reproducibility of the *Impaired decision-making* subscale was found to be poor.[27] Therefore, the total NWFQ score of the other 47 items was applied, with a scale range of 0-100.

Prior data of the same study population were used to define cut-off values for each subscale and for overall work functioning impairments,[27] based on percentiles in a group of workers with mental health complaints. The defined cut-off values selected 38% of the cases in that sample. In the design study,[17] we estimated this value to be 31%, which was not accurate.

Mental Health Complaints

Distress

Distress was measured with the distress subscale of the Four-Dimensional Symptoms Questionnaire (4DSQ).[28] Its 16 items had a total score range from 0-32. A cut-off point of ≥ 11 was applied.[29]

Need for recovery after work

The need for recovery after work (work-related fatigue) was measured using the Dutch Experience and Evaluation of Work questionnaire (Dutch: VBBA).[30] The 11-item subscale has a total scores ranging from 0-11. A cut-off point of ≥ 6 was applied (sensitivity: 72; specificity: 79).[31]

Risky drinking behavior

The 3-item AUDIT-C was used for risky drinking behavior.[32] Total scores ranged from 0-12. The cut-off score was ≥ 5 for men (sensitivity: 91; specificity: 68) and ≥ 4 for women (sensitivity: 92; specificity: 74).[33] It was discovered that male participants who scored above cut-off at baseline (N = 5) were not classified as such, due to a system error in the online questionnaire. Therefore, two men were not invited for the preventive consultation with their occupational physician as they did not screen positive on any other screener.

Depression and anxiety

Depression and anxiety were measured with the corresponding subscales of the Brief Symptom Inventory (BSI).[34] Each subscale has six items, and the total scores can range from 0-4. For both subscales, mean scores of ≥ 0.42 were used as cut-off points (for depression: sensitivity: 86; specificity: 66; for anxiety: sensitivity: 83; specificity: 62).[35]

Post-traumatic stress disorder

Post-traumatic stress disorder was measured using the Dutch translation of the Impact of Event Scale.[36] The scale includes 15 items; total scores range from 0-75. A cut-off point of ≥ 26 was applied.[37]

Analyses

Analyses were performed at the level of the individual worker, according to the intention-to-treat principle. As planned, the effect analyses on the primary and secondary outcome measures were performed only for the sample of subjects that were screened as positive for work functioning impairments and/or mental health complaints.

Dropout analysis

Possible differences of the study groups in effects on the primary outcome (help-seeking behavior) between completers and dropouts were first explored in a graph over time for completers and dropouts separately. If different patterns in effects were detected, significance was tested by Fisher's exact test. In case of statistically significant differences, multivariable logistic regression analysis was performed, with dropout as a dependent measure to identify potential predictive variables for dropout. In the case of significant effects in the logistic regression, the given variable was considered for inclusion as a covariate in the effect analyses.[38]

Analysis of confounders

All four possible confounders were tested for baseline differences between the study groups by Fishers' exact test. In the case of significant baseline differences, logistic regression analysis was performed to test for significant effects of these potential confounders on help-seeking behavior. If both these tests pointed to significant effects, the variable was included as a covariate in the effect analysis of help-seeking behavior.

Effect analysis

Generalized Linear Mixed Models (GLMM) were applied to analyze the differences between the two study groups in terms of dichotomous outcome measures (help-seeking behavior, intention to seek help, informal help-seeking behavior, and work as the focus of the consultation). Linear Mixed Models (LMM) were used to assess differences in continuous outcomes (impaired work functioning and all six mental health complaints) between the study groups. The continuous outcome measures with skewed distributions were log-transformed to meet the basic assumption of linear mixed models.

GLMM and LMM model construction

Outcomes at three and six months were included as dependent variables. Baseline scores were added as covariates in the model. As fixed effects, we included study group, time, and the interaction of study-group*time. Ward and ward*subject were included as random effects; however, if ward was not significant, it was excluded from the model. The effect of interest was the main effect of study-group and the interaction between study-group*time. The main effect of study-group could be interpreted as the differences between the two study groups over the follow-up period. The interaction term indicated a difference in time course on the outcome between the study groups. A p-value of $p < 0.05$ was considered statistically significant.

All analyses were performed using the statistical package PASW Statistics 18, except for the GLMM analyses, which were performed using the statistical package IBM SPSS Statistics 19.

RESULTS

Employee flow

Figure 1 presents the flow of participants. In total, 1,152 nurses and allied health professionals from 57 wards were invited. Twenty-nine wards with 591 workers were allocated to the WHS OP-care group and 28 wards with 561 workers were allocated to the control group. From March 15th until April 26th 2011, 421 workers were enrolled; 379 were eligible for participation in the study: 191 in the intervention and 188 in the control group. After six months, 99 workers in the intervention group and 126 workers in the control group had completed all three online questionnaires. In total, 129 workers were lost to follow-up, 79 in the intervention and 50 in the control group. The reasons for withdrawal from the study were not assessed.

In the WHS OP-care group, among the 151 positively screened workers, 51 workers (34%) accepted the invitation and consulted their occupational physician.

Participant characteristics at baseline

The characteristics of the participants are shown in Table 1. The majority of the participants were female, Dutch, and married. The mean age was 42 years, and approximately two-thirds of all participants were nurses.

Overall, 151 (79%) workers in the WHS OP-care group were screened as positive for work functioning impairment and/or mental health complaints and consequently received an invitation for a preventive consultation with their occupational physician. In the control group, 161 (86%) workers were indicated as either having work functioning impairments and/or mental health complaints and therefore were eligible for the analyses in this study.

In the GLMM and the LMM analyses, the random effect of ward was in no case statistically significant and therefore was dropped from the mixed model analyses.

Help-seeking behavior

Dropout analyses

A difference in help-seeking behavior at baseline between dropouts and completers in the control group was identified through the graphs: in the control group dropouts showed more help seeking at baseline than completers. This difference was statistically significant (Fisher's exact test ($p < 0.05$)). In a subsequent analysis, two variables were statistically significant predictors of dropout for the entire group: *experience with psychological or psychiatric care* ($p = 0.02$) and *intention to seek help from the occupational physician* ($p = 0.03$). Because no differences between the study groups were found at baseline, these factors were not included as covariates.

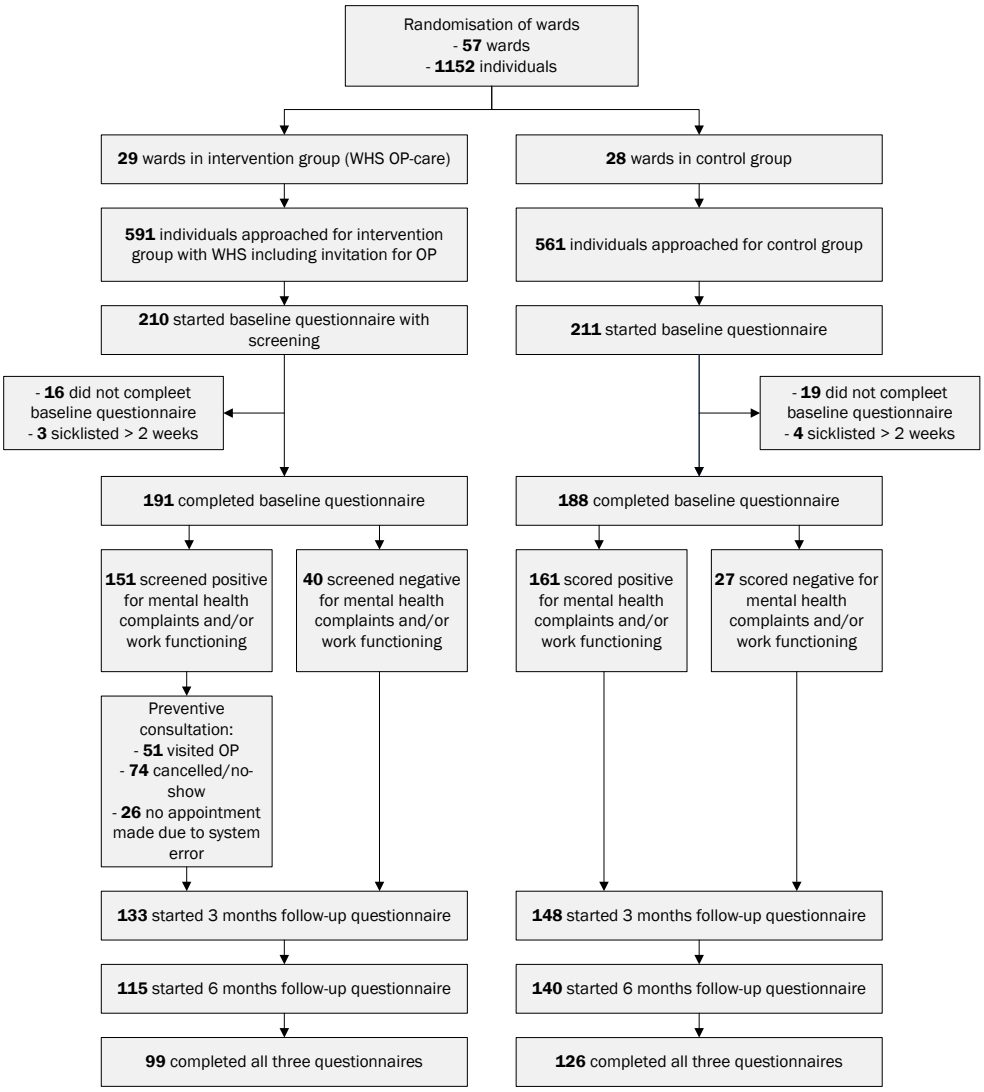


Figure 1. Flow chart of participants through the trial.

Table 1. Participant characteristics at baseline for the total sample and for the positively screened sample split for the WHS OP-care group and the control group.

	Total Sample				Positively screened sample			
	WHS OP-care (N=191)		Control (N=188)		WHS OP-care (N=151)		Control (N=161)	
Variable	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**
Gender								
Female	156	(82)	145	(77)	124	(82)	126	(78)
Age in years (mean (SD))	43	(11)	42	(11)	43	(11)	42	(12)
18-24	12	(6)	12	(6)	8	(5)	10	(6)
25-34	42	(22)	43	(23)	34	(23)	40	(25)
35-44	44	(23)	46	(25)	33	(22)	36	(22)
45-54	60	(31)	53	(28)	48	(32)	44	(27)
55-64	33	(17)	34	(18)	28	(19)	31	(19)
Marital status								
Married/living together with partner	131	(69)	123	(66)	100	(67)	100	(63)
Partner, not living together	12	(6)	18	(10)	10	(7)	16	(10)
Single	38	(20)	36	(19)	32	(21)	34	(21)
Divorced	6	(3)	9	(5)	5	(3)	9	(6)
Widowed/widower	3	(2)	1	(1)	3	(2)	1	(1)
Ethnic background								
Dutch	156	(83)	163	(87)	122	(82)	143	(89)
Immigrant	32	(17)	25	(13)	26	(18)	18	(11)
Occupation								
Nurse	116	(61)	134	(71)	89	(59)	115	(71)
Nurse practitioner	12	(6)	22	(12)	9	(6)	18	(11)
Surgical nurse	12	(6)	5	(3)	12	(8)	5	(3)
Anesthetic nurse	12	(6)	0	(0)	9	(6)	0	(0)
Allied health professional	39	(20)	27	(14)	32	(21)	23	(14)
Nursing specialization								
Yes	77	(66)	86	(64)	58	(65)	75	(65)
Years of experience (mean (SD))	12	(11)	11	(10)	13	(10)	11	(10)
≤ 1	17	(9)	21	(11)	14	(9)	18	(11)
2-5	50	(27)	49	(26)	35	(24)	42	(26)
6-10	33	(18)	39	(21)	25	(17)	35	(22)
11-15	30	(16)	25	(13)	27	(18)	21	(13)
16-20	14	(7)	17	(9)	10	(7)	14	(9)
21-25	14	(7)	13	(7)	13	(9)	9	(6)
>25	30	(16)	22	(12)	25	(17)	20	(13)
Working hours/week in contract (mean (SD))	29	(8)	31	(6)	29	(7)	31	(6)
Type of contract								
Permanent position	172	(92)	174	(93)	136	(92)	150	(94)
Fixed-term contract	9	(5)	12	(6)	7	(5)	9	(6)
Temporary employment	6	(3)	0	(0)	4	(3)	0	(0)
Other	1	(1)	1	(1)	1	(1)	1	(1)
Help-seeking behavior								
Sought help during past 3 months	115	(60)	117	(62)	98	(65)	102	(63)
Intention to seek help	9	(12)	9	(13)	6	(11)	9	(15)
Experience with mental health care	139	(74)	139	(74)	111	(74)	127	(79)
*Work functioning * (above cut-off)*								
Work functioning impairments	115	(60)	131	(70)	115	(76)	131	(81)
Mental health state (above cut-off)								
CMDs (above cut-off for ≥ CMDs)	112	(59)	119	(63)	112	(74)	119	(74)
Distress	50	(26)	48	(26)	50	(33)	48	(30)
Work related fatigue	52	(27)	65	(35)	52	(34)	65	(40)
Risky drinking behavior **	47	(25)	63	(34)	45	(30)	61	(38)
Depression	50	(26)	45	(24)	50	(33)	45	(28)
Anxiety	32	(17)	30	(16)	32	(21)	30	(19)
Post traumatic stress	21	(11)	19	(10)	21	(14)	19	(12)
Positively screened for CMD and/or impaired work functioning	151	(79)	161	(86)	151	(100)	161	(100)

* Work functioning is presented here including NWFQ subscale 2) *Impaired decision making*, as it was included in the baseline screening. ** Due to a technical error in the online questionnaire, male participants who scored above cut-off at baseline were not classified as such; the numbers presented here were corrected for this error.

Covariates analyses

No statistically significant baseline differences between the study groups were found for any of the four potential prognostic variables. These variables were therefore excluded from the analyses.

Effects on help-seeking behavior

Table 2 presents detailed descriptive data on help-seeking behavior. In the three months after baseline, among the group of workers that were screened as positive at baseline, 67% of workers in the WHS OP-care group and 54% in the control group had consulted one or more caregivers. The occupational physician was visited most often in the WHS OP-group (by 45% of the workers), whereas 6% of those in the control group visited their occupational physician. At the six-month follow-up, 48% of the workers in the WHS OP-care group showed help-seeking behavior, whereas 56% of those in the control group did.

A statistically significant interaction effect of study-group*time on help-seeking behavior ($p = 0.02$) was found (see Table 2). A trend main effect of study-group was found at the three-month follow-up ($p = 0.09$) but not at the six-month follow-up ($p = 0.14$). Therefore, we can conclude that a difference in course over time in help-seeking behavior was observed among the study groups. In Figure 2, it can be seen that in the WHS OP-care group help-seeking remained stable between baseline and three months follow-up, while it decreased for the control group (with a difference between the two groups of 13%). Towards six months follow-up, help-seeking behavior decreased more steeply in the WHS OP-care group than it did in the control group, resulting in even slightly less help-seeking behavior (8% lower).

Secondary outcomes on work functioning and mental health complaints

Descriptive analyses for work functioning and mental health complaints at all three measurement points are presented in Table 3.

Work functioning impairments

A significant effect between groups was found for work functioning ($p = 0.04$). Workers in the WHS OP-care group showed statistically significant less work functioning impairments over time than workers in the control group. The mean difference between the study groups was 1.5 points for the total follow-up period, based on the estimated means of the LMM analyses. The difference between the study groups seems small, although statistically significant. Therefore, for more insight into this effect, the relative change scores of individual workers were compared by applying the minimal important change (MIC) values of the NWFQ total scale (40% improvement). In the WHS OP-care group, 45% of the workers showed minimal important improvements or more on the NWFQ total score compared to 30% of the workers in the control group ($p = 0.03$). At six months of follow-up, 41% of workers in the WHS OP-care had improved importantly or more compared to 28% in the control group ($p = 0.05$).

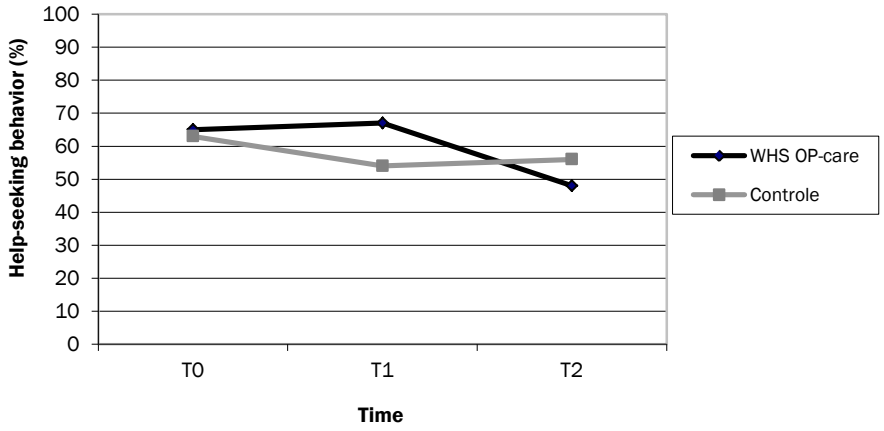


Figure 2. Help-seeking behavior in percentages of workers for the two study groups at the three measurement points.

Mental health complaints

Risky drinking behavior was the only mental health complaint for which a statistically significant effect was found. The interaction of study-group*time was statistically significant ($p < 0.01$). However, the difference in scores on risky drinking behavior between the study groups was not statistically significant at three months of follow-up ($p = 0.08$) or at six months of follow-up ($p = 0.36$). Thus, we can conclude that a difference in course over time in risky drinking behavior was observed among the study groups.

DISCUSSION

In total, 191 nurses and allied health professionals took part in a WHS mental module. In the WHS OP-care group, 151 (79%) of the workers were classified as having work functioning impairments and/or mental health complaints compared to 161 (86%) in the control group.

The hypothesis of an increase in help-seeking behavior for the WHS OP-care group was confirmed. A statistically significant interaction effect was found, thus, the time course of help-seeking behavior differed significantly for the two study groups. Also, a marginal significant difference between the two study groups at the three-month follow-up was found. These effects are based on an increase in help seeking at the occupational physician only.

Table 2. Descriptives on primary and secondary outcome measures on help-seeking behavior at baseline, three months of follow-up, and six months of follow-up and number and type of caregivers visited; p-value of Generalized Linear Mixed Model analyses on three and six months of follow-up.

Outcome measures	Baseline		3 months		6 months		P-value (GLMM analyses)	
	Relative frequency	(%)	Relative frequency	(%)	Relative frequency	(%)	Study-group	Study-group*time
Help-seeking behavior for all caregivers								
WHS OP-care	98/151	(65)	68/101	(67)	42/87	(48)	0.907	**0.016**
Control	102/161	(63)	68/127	(54)	66/117	(56)		
Intention to seek help								
WHS OP-care	6/53	(11)	8/42	(19)	5/45	(11)	0.220	0.794
Control	9/59	(15)	13/59	(22)	13/52	(25)		
Informal help seeking								
WHS OP-care	49/151	(33)	31/101	(31)	19/87	(22)	0.722	0.366
Control	60/161	(37)	33/127	(26)	27/117	(23)		
Work as content of help (OP and supervisor excluded)								
WHS OP-care	35/82	(43)	19/45	(42)	13/33	(39)	0.071	0.341
Control	36/80	(45)	25/55	(46)	25/59	(42)		
Number of different caregivers visited in past 3 months								
0 caregivers								
WHS OP-care	53/151	(35)	33/101	(33)	45/87	(52)		
Control	59/161	(37)	59/127	(47)	51/117	(44)		
1 or 2 caregivers								
WHS OP-care	72/151	(48)	48/101	(48)	35/87	(40)		
Control	83/161	(52)	54/127	(43)	46/117	(39)		
> 2 caregivers								
WHS OP-care	26/151	(17)	20/101	(20)	7/87	(8)		
Control	19/161	(12)	14/127	(11)	20/117	(17)		

Table 2. (continued)

Outcome measures	Baseline		3 months		6 months		P-value (GLMM analyses)	
	Relative frequency	(%)	Relative frequency	(%)	Relative frequency	(%)	Study-group	Study-group*time
Help-seeking behavior per caregiver in past 3 months								
General Practitioner								
WHS OP-care	58/151	(38)	31/101	(31)	22/87	(25)		
Control	56/161	(35)	38/127	(30)	43/117	(37)		
Supervisor								
WHS OP-care	56/151	(37)	24/101	(24)	19/87	(22)		
Control	55/161	(34)	35/127	(28)	36/117	(31)		
Occupational physician								
WHS OP-care	20/151	(13)	45/101	(45)	4/87	(5)		
Control	13/161	(8)	7/127	(6)	11/117	(9)		
Physiotherapist								
WHS OP-care	26/151	(17)	15/101	(15)	8/87	(9)		
Control	16/161	(10)	14/127	(11)	12/117	(10)		
Psychologist, psychiatrist and/or coach								
WHS OP-care	15/151	(10)	11/101	(11)	7/87	(8)		
Control	22/161	(14)	19/127	(15)	18/117	(15)		
Social worker (in company) and/or religious counselor								
WHS OP-care	3/151	(2)	1/101	(1)	0/87	(0)		
Control	4/161	(3)	3/127	(2)	3/117	(3)		
Alternative treatments								
WHS OP-care	8/151	(5)	4/101	(4)	5/87	(6)		
Control	9/161	(6)	6/127	(5)	7/117	(6)		

Table 3. Descriptives on secondary outcome measures impaired work functioning and mental health complaints at baseline, three months of follow-up, and six months of follow-up; p-values of Linear Mixed Model analyses on three and six months of follow-up.

Outcome measures	Baseline				Three months				Six months				P-value (LMM analyses)	
	Above cut-off		Median (range)		Above cut-off		Median (range)		Above cut-off		Median (range)			
	Relative frequency	(%)			Relative frequency	(%)			Relative frequency	(%)			Study-group	Study-group* time
Work functioning (NWFQ 0-100)*														
WHS OP-care	88/151	(58)	12	(0-51)	38/100	(38)	9	(0-38)	34/86	(40)	8	(0-62)	**0.039**	0.842
Control	110/161	(68)	14	(0-54)	61/124	(49)	10	(0-38)	60/116	(52)	10	(0-44)		
Mental health complaints														
Distress (4DKL, 0-32)														
WHS OP-care	50/151	(33)	6	(0-30)	17/99	(17)	4	(0-28)	14/86	(16)	5	(0-31)	0.243	0.444
Control	48/161	(30)	6	(0-32)	25/123	(20)	5	(0-29)	26/116	(22)	5	(0-30)		
Need for recovery (VBBA, 0-100)														
WHS OP-care	52/151	(34)	27	(0-100)	30/98	(31)	18	(0-91)	28/85	(33)	27	(0-100)	0.358	0.697
Control	65/161	(40)	36	(0-100)	42/123	(34)	27	(0-100)	39/116	(34)	36	(0-100)		
Risky Drinking (AUDIT-C, 0-12) **														
WHS OP-care	45/151	(30)	3	(0-10)	30/98	(31)	3	(0-7)	21/85	(25)	2	(0-6)	0.666	**0.004**
Control	61/161	(38)	3	(0-9)	35/123	(29)	3	(0-8)	39/116	(34)	3	(0-8)		
Depression (BSI, 0-4)														
WHS OP-care	50/151	(33)	0.17	(0-2.17)	22/97	(23)	0.17	(0-3)	21/85	(25)	0.17	(0-4)	0.405	0.624
Control	45/161	(28)	0.17	(0-3.33)	35/123	(29)	0.17	(0-1.5)	30/116	(26)	0.17	(0-3.33)		
Anxiety (BSI, 0-4)														
WHS OP-care	32/151	(21)	0.17	(0-1.67)	14/97	(14)	0	(0-1)	10/85	(12)	0	(0-3.5)	0.589	0.943
Control	30/161	(19)	0.17	(0-2.83)	13/122	(11)	0.17	(0-2.5)	22/116	(19)	0.17	(0-2.67)		
PTSS (SVL, 0-75)														
WHS OP-care	21/151	(14)	2	(0-75)	7/97	(7)	0	(0-47)	7/85	(8)	0	(0-65)	0.828	0.702
Control	19/161	(12)	3	(0-48)	13/122	(11)	0	(0-62)	9/116	(8)	0	(0-48)		

* NWFQ total scores are calculated without without subscale 2) *Impaired decision making*.

** Due to a technical error in the online questionnaire, male participants who scored above cut-off at baseline were not classified as such; the numbers presented here were corrected for this error.

The hypothesized larger decrease of work functioning impairments in the WHS OP-care group was confirmed. The hypothesis of an improvement of mental health complaints in the WHS OP-care group could only partly be confirmed, as only for risky drinking behavior an interaction effect was observed.

Methodological considerations

The design with cluster randomization and pre-randomization was a strength of this study. Randomization at the ward level reduces the contamination of study groups. Furthermore, the pre-randomization approach allows the participants to be blinded with respect to the information related to the other study groups. No indication of a contamination effect appeared during the study.

One limitation is the lack of a supplementary per-protocol analysis, in which effects would be analyzed separately for the group of workers that followed the invitation of the occupational physician. This extra information might have helped in the interpretation of any effects found on work functioning and/or mental health state.

A second methodological issue concerns the power. For sufficient power, our sample size calculation resulted in a required 114 workers per condition that were screened as positive and therefore were eligible for inclusion in the effect analyzes. This sample size was not reached. As a consequence, the probability for type II errors increases, which is the risk of non-significant p-values despite trends for differences between the study groups. This might have occurred regarding the main effect of study-group on help-seeking behavior at three months of follow-up.

Interpretation

Notably, the percentage of workers classified as having work functioning impairments and/or mental health complaints at baseline was high (approximately 80%) in this study. Not all workers who were screened as positive can be seen as workers with CMDs or serious work functioning problems. However, all of these workers had reported symptoms of either work functioning impairments or a mental health complaint, which were at least classifiable as mild symptoms.

This high percentage can be partly explained by the choice of cut-off values for the screeners. Cut-off values were selected with optimal sensitivity at the cost of specificity, as it was desired not to miss the chance to offer workers the opportunity for a preventive consultation with the occupational physician. All cut-off values for the mental health screeners applied in this study were validated, and all the mental health complaints screened for are known to be related to the work of nurses and allied health professionals. We assume that the high number of screeners (seven for work functioning and six for mental health complaints) resulted in the high overall percentage of positively screened workers. It is important to note that the applied screening method was regarded as safe and did not affect the workers

negatively for two reasons. First, the personal feedback on screening results that was provided online, was formulated mildly without speaking of *health problems* or *diagnosis* to prevent any incorrect interpretation or premature conclusions. Second, the consultation with the occupational physician can be regarded as a second screening stage; no treatment was initiated before the visit to the occupational physician (or any other caregiver).

Overall, we assume that the screening strategy was suitable as a first step for screening workers for whom a preventive consultation with the occupational physician was regarded as useful. However, in the future, more cautious screening with cut-off values that allow for higher specificity might be advisable when providing feedback online, as the positive predictive value will generally increase.

The WHS mental module affected the help-seeking behavior of workers who received the WHS mental module positively in the first three months of follow-up. However, this effect is based on more visits to the occupational physician only. It might be possible that visiting the occupational physician for a preventive consultation led to a decreased need for help later on, as we observed less help-seeking behavior and a lower intention to seek help at the six-month follow-up among the WHS OP-care group (although this tendency was not statistically significant). Overall, we assume that the preventive aim of the WHS mental module in stimulating help-seeking behavior in workers with indications for work functioning impairments and/or mental health complaints was achieved.

Additionally, the WHS mental module led to the expected decrease in work functioning impairments with considerable differences in the number of workers that improved above the minimal important change value. Thus, the increase in help-seeking behavior, although based on the number of initiated visits to the occupational physician, seems to be effective for one part of the problem screened for: work functioning. For the mental health complaints, however, no effects were found except for risky drinking behavior. Therefore, we might conclude that the WHS mental module has an effect on behavioral outcomes, i.e., work functioning and risky drinking behavior, whereas not on the level of mental health complaints. Based on the process evaluation (Ketelaar et al., submitted), we assume that the occupational physicians might have provided advice only to a small extent, as they expressed some reluctance to initiate interventions for workers in this preventive setting. However, in this study the type and extent of advice provided was not assessed in detail.

In previous studies that combined screening for mental health and work aspects with subsequent interventions, considerable effects on depressive symptoms[39] [40] and on work functioning[41] were found. In all three studies, the treatment offered was much more intensive, compared to our study, but was also offered to workers with more serious mental health complaints. More intensive guidance for a small selected group of workers with moderate severe to severe mental health complaints might also be useful within the WHS mental module, possibly after referral to specialized mental caregivers, such as psychologists or

welfare workers. For the group of workers with milder symptoms, less intensive interventions are regarded as sufficient, which the occupational physician should offer more proactively. Any advice given by the occupational physician should aim at preventing the aggravation of symptoms and, if necessary, initiating temporary work adaptations to guarantee the safety of the workers and patients. To ensure more proactive guidance by the occupational physician, the training of the WHS mental module should focus on possible reluctance and should intensively discuss communication strategies for workers with mild mental health complaints but without an expressed need for help. With these adaptations, the effects of the WHS mental module on mental health might be elucidated. Furthermore, the effects on work functioning and risky drinking behavior might be enhanced.

Implications for practice

The WHS mental module studied was rather innovative in the fact that feedback on screening results was provided online. For the future, we recommend that the choice for online feedback and the applied screeners should be more finely tuned. When workers are provided with feedback personally by the occupational physicians, the choice for high sensitivity but low specificity is tolerable. When choosing to present personal screening results online, cut-off values with higher specificity and lower sensitivity should be chosen.

The external validity is regarded to be high for any other hospital setting, as the study was executed in a real-life setting.

Recommendations for future research

The effect the WHS mental module has on work functioning is promising; however, little is known about the type of advices or interventions that occupational physicians provided to workers during the preventive consultation and the adaptations the workers initiated in their work or behavior. With more insight into the work functioning interventions, the preventive consultation protocol could be improved. This might lead to an enhanced effect on work functioning.

Conclusion

We conclude that the WHS mental module is a successful strategy to stimulate seeking help from the occupational physician. Regarding the aims of WHS as formulated by the Netherlands Society of Occupational Medicine,[22] we can conclude that the WHS mental module studied here does monitor and improve work functioning. However, the aim to monitor and improve mental health was not met sufficiently by the studied WHS mental module over the six months follow-up period.

REFERENCES

1 Wang J, Adair CE, Patten SB. Mental health and related disability among workers: a population-based study. *Am J Ind Med* 2006;49:514-22.
2 Bültmann U, Rugulies R, Lund T et al. Depressive symptoms and the risk of long-term sickness absence: a prospective study among 4747 employees in Denmark. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:875-80.
3 Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401-10.
4 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
5 Altamura AC, Dell'Osso B, Vismara S et al. May duration of untreated illness influence the long-term course of major depressive disorder? *Eur Psychiatry* 2008;23:92-6.
6 Dahlberg KM, Waern M, Runeson B. Mental health literacy and attitudes in a Swedish community sample - investigating the role of personal experience of mental health care. *BMC Public Health* 2008;8:8.
7 Lexis MA, Jansen NW, Stevens FC et al. Experience of health complaints and help seeking behavior in employees screened for depressive complaints and risk of future sickness absence. *J Occup Rehabil* 2010;20:537-46.
8 Codony M, Alonso J, Almansa J et al. Perceived need for mental health care and service use among adults in Western Europe: results of the ESEMeD project. *Psychiatr Serv* 2009;60:1051-8.
9 Sareen J, Jagdeo A, Cox BJ et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. *Psychiatr Serv* 2007;58:357-64.
10 Adams EF, Lee AJ, Pritchard CW et al. What stops us from healing the healers: a survey of help-seeking behaviour, stigmatisation and depression within the medical profession. *Int J Soc Psychiatry* 2010;56:359-70.
11 Putnik K, de JA, Verdonk P. Road to help-seeking among (dedicated) human service professionals with burnout. *Patient Educ Couns* 2011;83:49-54.
12 Joyce T, McMillan M, Hazelton M. The workplace and nurses with a mental illness. *Int J Ment Health Nurs* 2009;18:391-7.
13 International Labour Organization (ILO). Technical and ethical guidelines for workers' health surveillance. Geneva: ILO, 1998.
14 Koh D, Aw TC. Surveillance in occupational health. *Occup Environ Med* 2003;60:705-10, 633.
15 Netherlands Society of Occupational Medicine (NVAB). Guidance Workers' Health Surveillance [In Dutch: Leidraad Preventief Medisch Onderzoek]. Utrecht, the Netherlands: NVAB, 2005.
16 Plat MJ, Frings-Dresen MH, Sluiter JK. A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *Int Arch Occup Environ Health* 2011;84:839-57.
17 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.
18 Schellings R, Kessels AG, ter RG et al. Indications and requirements for the use of prerandomization. *J Clin Epidemiol* 2009;62:393-9.
19 van der Klink JJ. Guideline: management of mental health problems of workers by occupational physicians [In Dutch: Richtlijn: Handelen van de bedrijfsarts bij werkenden met psychisch problemen]. Utrecht, the Netherlands: NVAB [Netherlands Society of Occupational Medicine], 2007.
20 Walters K, Buszewicz M, Weich S et al. Help-seeking preferences for psychological distress in primary care: effect of current mental state. *Br J Gen Pract* 2008;58:694-8.
21 Wilson CJ, Deane FP, Ciarrochi J et al. Measuring Help-Seeking Intentions: Properties of the General Help-Seeking Questionnaire. *Can J Counsell* 2005;39:15-28.
22 Judd F, Komiti A, Jackson H. How does being female assist help-seeking for mental health problems? *Aust N Z J Psychiatry* 2008;42:24-9.
23 Gudmundsdottir G, Vilhjalmsson R. Group differences in outpatient help-seeking for psychological distress: results from a national prospective study of Icelanders. *Scand J Public Health* 2010;38:160-7.
24 Rudell K, Bhui K, Priebe S. Do 'alternative' help-seeking strategies affect primary care service use? A survey of help-seeking for mental distress. *BMC Public Health* 2008;8:207.
25 Schomerus G, Matschinger H, Angermeyer MC. The stigma of psychiatric treatment and help-seeking intentions for depression. *Eur Arch Psychiatry Clin Neurosci* 2009;259:298-306.
26 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. Impaired work functioning due to common mental disorders in nurses and allied health professionals: the Nurses Work Functioning Questionnaire. *Int Arch Occup Environ Health* 2011;85:125-38.

27 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. Psychometric Properties of the Nurses Work Functioning Questionnaire (NWFQ). *PLoS One* 2011;6:e26565.
28 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.
29 van Rhenen W., van Dijk FJ, Schaufeli WB et al. Distress or no distress, that's the question: A cutoff point for distress in a working population. *J Occup Med Toxicol* 2008;3:3.
30 van Veldhoven M, Meijman TF. Het meten van een psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van arbeid (VBBA). Amsterdam: NIA, 1994.
31 Broersen JPJ, Fortuin RJ, Dijkstra L et al. Monitor Arboconvenanten: kengetallen en grenswaarden. *TBV* 2004;12:100-4.
32 Bush K, Kivlahan DR, McDonell MB et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998;158:1789-95.
33 Gual A, Segura L, Contel M et al. Audit-3 and audit-4: effectiveness of two short forms of the alcohol use disorders identification test. *Alcohol Alcohol* 2002;37:591-6.
34 de Beurs E. Brief Symptom Inventory (BSI): Manual [In Dutch: Brief Symptom Inventory (BSI): Handleiding]. Leiden: Pits Publishers, 2006.
35 de Beurs E, Zitman FG. The Brief Symptom Inventory (BSI): The reliability and validity of a brief alternative of the SCL-90 [In Dutch: De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90]. *Maandblad Geestelijke Volksgezondheid* 2006;61:120-41.
36 Brom D, Kleber RJ. The Impact of Event Scale [In Dutch: De Schok Verwerkings Lijst]. *Nederlands Tijdschrift voor de Psychologie* 2011;40:164-8.
37 Chemtob CM, Tomas S, Law W et al. Postdisaster psychosocial intervention: a field study of the impact of debriefing on psychological distress. *Am J Psychiatry* 1997;154:415-7.
38 DeSouza CM, Legedza AT, Sankoh AJ. An overview of practical approaches for handling missing data in clinical trials. *J Biopharm Stat* 2009;19:1055-73.
39 Lexis MA, Jansen NW, Huibers MJ et al. Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial. *Occup Environ Med* 2011;68:400-7.
40 Wang PS, Simon GE, Avorn J et al. Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. *JAMA* 2007;298:1401-11.
41 Lerner D, Adler D, Hermann RC et al. Impact of a Work-Focused Intervention on the Productivity and Symptoms of Employees With Depression. *J Occup Environ Med* 2012;54:128-35.

8

Mental Vitality @ Work – A workers' health surveillance mental module for nurses and allied health professionals: process evaluation of a randomized-controlled trial

Sarah M. Ketelaar
Fania R. Gärtner
Linda Bolier
Odile Smeets
Karen Nieuwenhuijsen
Judith K. Sluiter

Submitted for publication.

ABSTRACT

Background: To monitor and improve work functioning and mental health of nurses and allied health professionals, a workers' health surveillance (WHS) mental module was developed. A randomized controlled trial compared two strategies: an occupational physician care (OP-care) group and an e-mental health care (EMH-care) group. Both groups received screening and personalized feedback. When screened as positive, participants received either an invitation for a preventive consultation with an occupational physician (OP), or an invitation to start e-mental health intervention(s). This study aimed to assess participants' response and compliance to the intervention and OPs' adherence to the consultation protocol, and to describe the participants' and OPs' perspective on the WHS mental module.

Methods: All nurses and allied health professionals working in one Dutch academic medical center were invited. Process indicators were assessed using the occupational healthcare service's (OHS) recordings, self-report questions at three and six months follow-up, track & trace data, an OP consultation checklist, and a questionnaire and group interview with participating OPs. Numerical data were analyzed using descriptive statistics.

Results: The response of employees was 32%. Of the invited participants in the OP-care group, 41% went to the preventive consultation. Almost all of them felt they could be open and honest with their OP, and 70-80% felt their personal results reflected their work functioning and mental health. The preventive consultation was perceived effective by nine out of 15 participants. The OPs were satisfied with the consultation protocol and training, and their adherence to the consultation protocol was high. OPs regarded the preventive consultations, as well as implementing a WHS mental module in the future to be meaningful. Of participants in the EMH-care group, 17% logged into an e-mental health intervention. Most participants would appreciate to be offered a WHS in the future.

Conclusions: The WHS mental module was well-received by the study participants, and it is possible to include it in the OHS' activities. However, response and compliance should be improved. Analyses of the (cost-)effectiveness of the interventions will determine whether they were sufficient to improve help-seeking behavior, work functioning, mental health and well-being.

Trial registration: NTR2786

BACKGROUND

Working in healthcare places high mental demands on employees. Nurses, the largest occupational group in hospitals, are at higher risk of developing mental health problems compared to other human service workers and other healthcare workers.[1] Not only does their work pose a threat to their mental health, but suffering from mental health complaints also impairs their work functioning. Gärtner and colleagues[2] showed that when nurses have impaired mental health, impairments in various aspects of their work functioning may also exist, including task-related, intrapersonal and interpersonal aspects. Furthermore, impairments in work functioning can have serious effects for the nurses themselves but also for patients and their safety. Adding to this, the workforce is ageing, increasing the importance of sustainable labor participation. Consequently, it is essential to keep nurses mentally fit so they can continue to meet the high mental demands of modern-day work.

In the Netherlands, the occupational healthcare that is provided for employees with mental health problems can be considered effective. Care according to the occupational physicians' guidelines for treating employees who are absent from work due to impaired mental health has been demonstrated to improve mental health and to enhance return to work.[3 4] However, employees often do not seek help for their mental health complaints, and the occupational health service in the Netherlands often does not come into view until the employee is sick-listed. To ensure that early help can be provided before impaired work functioning and mental health problems worsen to the extent that employees cause serious incidents or call in sick, preventive actions are needed. Early identification of impaired work functioning and mental health complaints to provide timely help may be a first step in the prevention of more serious consequences for the mental health and safety of nurses and their patients.

One possible strategy for the early identification of impaired work functioning and mental health complaints is to periodically offer employees a mental module of a workers' health surveillance (WHS). WHS is an important component in the practice of occupational healthcare. It aims to detect negative health effects of work at an early stage to enable timely interventions.[5] Legislation in the Netherlands dictates that employers should periodically enable employees to undergo such an examination which is aimed at preventing or reducing the risks that the work poses for the employees' health.

Little is known about WHS targeting mental health effects.[6] With a WHS mental module, help-seeking behavior might be stimulated and work functioning and mental health might be improved. The Mental Vitality @ Work study was developed to examine the effects of a job-specific WHS mental module for nurses and allied health professionals on help-seeking behavior, work functioning, mental health and well-being by means of a cluster randomized controlled trial design. The WHS mental module combined screening with interventions. Two different strategies were compared: a strategy in which an occupational physician was

involved, and an e-mental health care strategy in which a tailored choice of online self-help mental health interventions was offered immediately after online screening. The effects of these two strategies on the study outcomes are presented in a separate article.

A process evaluation was performed alongside the Mental Vitality @ Work study to offer insight into the 'black box' of the intervention and to facilitate interpretation of the outcome results.[7-9] A process evaluation can shed light on the extent to which the intervention was applied to the targeted population as intended (e.g. compliance, adherence, satisfaction). Furthermore, it can offer insight into the factors of success and failure of the intervention and thus provide possible targets to improve the intervention.[10] To gain insight into the two interventions that were compared in the Mental Vitality @ Work study, the following aims were formulated for the process evaluation:

1) to describe the participants' response and compliance to the intervention,
2) to describe the participants' perspective on the WHS mental module,
3) to describe the occupational physicians' adherence to the consultation protocol, and
4) to describe the occupational physicians' perspective on the WHS mental module.

METHODS

Study design

The Mental Vitality @ Work study was designed as a cluster randomized controlled trial with block randomization carried out at ward level.[11] The complete trial, described in Gärtner and colleagues,[11] included two intervention groups and one control group. This process evaluation focuses on the two intervention groups: a) the occupational physician care (OP-care) group, who received screening, personal results and an invitation for a preventive consultation with an occupational physician, and b) the e-mental health care (EMH-care) group, who received screening, personal results and an invitation to follow an e-mental health intervention.

Participants

The study population of the complete trial consisted of all nurses, including surgical nurses and anesthetic nurses, and allied health professionals working in one academic medical center in the Netherlands. All eligible employees were invited to take part in the study, which for the complete trial added up to 1731 employees working in 86 wards. After randomization, 591 employees were assigned to the OP-care group and 579 employees to the EMH-care group, together working in 57 wards. Participants were included if they were not, or were not expecting to be on sick leave for more than two weeks at the start of the study.

Procedure

At the start of the study (March 2011), the potential participants received an invitation in their work e-mail account to take part in the study. It was possible to discontinue the questionnaire and complete it after logging in again, within six weeks.

After completing the baseline questionnaire, participants immediately received their personal results electronically as well as the intervention specific to their study group. The first follow-up questionnaire after three months and the second follow-up questionnaire after six months were sent to baseline participants.

Recruitment strategy

The board of management of the academic medical center and the works council were informed about all details of the study and gave their consent. An advisory committee was formed with the head of Human Resources, the head of the in-house occupational health service (OHS), a nursing director, and a director of a ward with allied health professionals. The works council and the advisory committee were regularly informed about the progress of the study.

To inform all relevant parties within the academic medical center, the following activities were carried out: information was published in the employee magazine; posters were hung in the staff rooms of each ward and in public places, such as outside the restaurant, in the halls and squares, beside the locker rooms, and beside the elevators; a banner was shown twice on the intranet with a link to information about the study; and the researchers stood in public places with flyers and small gifts such as fruit, little water bottles, and chocolates.

Before the start of the study and each follow-up questionnaire, the nursing directors, head nurses and managers of allied health professionals were sent information about the study. At three months follow-up, frequently asked questions were addressed in this information. The head nurses and managers of allied health professionals were also sent information which they could use to inform their employees. Furthermore, head nurses and managers of eligible participants who would receive personal results and an invitation for an intervention were telephoned to answer any questions they had about the study.

An information letter was sent to the home address of eligible participants, to remind them to open their work e-mail account and to motivate them to read the invitation to fill out the questionnaire. For three and six months follow-up, a lottery (in which a voucher for a weekend in a wellness hotel would be raffled) was announced in this letter. The researchers placed flyers and small gifts in the staff rooms of each ward twice during the six weeks that the questionnaires could be filled out. Furthermore, three reminders were sent by e-mail to those who had not yet participated or had not finished filling out the questionnaire. In these reminders, frequently asked questions were addressed.

Computer rooms were reserved at certain times to facilitate filling out the questionnaire, which was advertised in the invitation e-mail and the reminders. As very few people came to

these computer rooms, it was decided not to continue this after the baseline procedure. If participants telephoned for help with logging into or filling out the online questionnaire, the researchers assisted them by phone or went to the participants' workplace if this was needed.

Interventions

The OP-care intervention and the EMH-care intervention both consisted of online screening and immediate feedback on personal results on work functioning and mental health outcomes, followed by a group-specific intervention.

Screening and personal results

At baseline, participants were screened for the following aspects (for details see Gärtner and colleagues[11]): impaired work functioning, distress, work related fatigue, risky drinking behavior, depression (including suicide risk), anxiety (including panic disorder), and posttraumatic stress disorder. The work relatedness of any mental health complaints was also assessed. The participants received feedback on their personal screening results in two ways: onscreen immediately after filling out the baseline questionnaire, and in an e-mail sent to the address they provided to this end.

Occupational physician care

For participants in the OP-care group who screened positive on impaired work functioning and/or on any mental health complaints, personal results were followed by an invitation for a face-to-face preventive consultation with their own occupational physician (OP), within two weeks. They were then sent a letter to their home address with an appointment for this consultation. The consultation was voluntary and supervisors were not informed about it.

A consultation protocol was written which was based on interviews with the participating OPs to elicit their current practice and on the evidence-based guideline for OPs' treatment of employees with mental health problems.[4 12] It consisted of seven steps: 1) discussing expectations; 2) discussing screening results and characteristics of work functioning/mental health complaints; 3) discussing possible causes in the private, work and health condition and consequences for work functioning; 4) identifying the problem and offering rationale; 5) giving advice on how to tackle health complaints, how to improve work functioning, how to prevent consequences of impaired work functioning, and how to communicate with the supervisor about work functioning and mental health; 6) discussing possible follow-up or referral to other care providers; and 7) summarizing the consultation. All participating OPs received a three-hour training in using the protocol for the consultations.

E-mental health care

For participants in the EMH-care group, the personal results were followed by a tailored choice of self-help e-mental health interventions. Participants who screened positive on one or more

mental health aspects were invited to follow one or more e-mental health interventions. Participants who screened negative on all mental health aspects were invited to follow an e-mental health intervention aimed at enhancing and retaining mental fitness. The e-mental health interventions were developed as stand-alone interventions by the Trimbos Institute (Netherlands Institute of Mental Health and Addiction) at an earlier stage. The e-mental health interventions were:

- *Psyfit*:[13] aimed at enhancing mental fitness;
- *Strong at work*: aimed at gaining insight into work stress and learning skills to cope with work stress;
- *Colour your Life*:[14-19] aimed at tackling depressive symptoms;
- *Don't Panic Online*:[20-24] aimed at reducing panic symptoms for subclinical and mild cases of panic disorder;
- *Drinking Less*:[25 26] aimed at reducing risky drinking behavior.

To tailor the specific choice of e-mental health interventions to the individual's work functioning and mental health state, an algorithm was developed by the researchers. For example, if screened positive on (partly) work related depression, participants were advised to choose one of the following interventions: *Colour your Life, Strong at Work* or *Psyfit*.

In case of positive screening on work functioning, participants received an onscreen educational leaflet on how to improve their work functioning.

Data collection

Several methods were used to collect data for the process evaluation. Table 1 offers a schematic presentation of the process evaluation elements and the methods used to measure these elements.

Participants' response and compliance to the intervention

To determine response, the researchers registered the number of participants that took part in the study. In the OP-care group, five process indicators (prescripted answers) were asked to measure compliance (received invitation for preventive consultation, went to consultation,

Table 1. Elements of the process evaluation and the methods used to measure each element.

Elements	Measurement							
	Regist-ration resear-chers	Regist-ration OHS	Checklist OPs	Track & trace EMH	Quest. 3 months	Quest. 6 months	Quest. OPs	Inter-view OPs
Response participants	X							
Compliance participants		X		X	X			
Perspective participants			X		X	X		
Adherence OPs		X	X					
Perspective OPs							X	X

Whether participants were sent an invitation for a preventive consultation and whether they came to the appointment was recorded by the OHS.

Compliance in the EMH-care group was assessed with four indicators (prescripted answers) (followed e-mental health interventions, received onscreen educational leaflet on how to improve work functioning, followed the advice in this leaflet, and discussed results and/or advice with supervisor), and reasons were asked for choosing a specific e-mental health intervention and for not following an e-mental health intervention (open-ended). Participants could also explain why they did or did not follow the advice on how to improve their work functioning (open-ended).

Participants who logged into an e-mental health intervention were followed using 'track & trace', which means that the intervention automatically kept track of the activities of the participant in the e-mental health intervention. Whether participants logged into an e-mental health intervention at least once and the number of modules on which they started were used as indicators.

Participants' perspective on the WHS mental module

The participants' perspective on the WHS mental module encompassed four aspects. Their perspective on the preventive consultation and the usefulness of the advice that was given (OP-care group, prescripted answers) and on the e-mental health interventions (EMH-care group, graded) was asked. Secondly, it was assessed whether they would want to receive their personal results (both groups) and an invitation for their OP differently in future (OP-care group) (prescripted answers). Regarding perceived effectiveness, the participants' perception of the effectiveness of the advice that was given by the OP (OP-care group) and of following the e-mental health interventions and the advice on how to improve work functioning (EMH-care group) was asked (prescripted answers). Participants in both groups were asked about their preferences regarding future implementation of the WHS mental module (prescripted answers).

For participants in the OP-care group who came to the preventive consultation, the OPs recorded whether the participant felt that their screening results indeed reflected their mental health condition and work functioning (*completely/partially/not*).

Occupational physicians' protocol adherence

The OPs filled out a checklist for each preventive consultation they held, containing four questions with prescripted answers (completion of each step of the consultation protocol, advice given to the employee, referral of the employee, and scheduled follow-up). In addition, the OHS recorded how many follow-up consultations the participants went to in the six months after baseline.

Occupational physicians' perspective on the WHS mental module

The OPs' perspective encompassed their perspective on the WHS mental module as performed in the study, and their perspective and preferences regarding future implementation of a WHS mental module. These aspects were measured by means of a short questionnaire and a semi-structured group interview, after all the preventive consultations had taken place.

Statistical analyses

The analyses were performed using the statistical package PASW Statistics 18 and Microsoft Office Excel 2007. Numerical data were analyzed using descriptive statistics. To facilitate interpretation of open-ended questions, answers that were similar were clustered. For each cluster, the proportion was calculated of the total number of answers on that question. For the results of the semi-structured group interview with the participating OPs, a summary was given per question and presented as a written description.

RESULTS

Participants' response

In total, 191 participants (32%) in the OP-care group and 178 (31%) in the EMH-care group finished the baseline questionnaire and were included in the analyses. Baseline characteristics are shown in Table 2. Respectively 82% and 83% of the participants were female. The average age of the participants was 43 years in the OP-care group and 37 years in the EMH-care group. In both groups, nurses constituted the largest number of participants (61% versus 73%).

In the OP-care group, 151 (79%) scored above any of the cut-off points of impaired work functioning and/or impaired mental health and were eligible for a preventive consultation. Due to a technical and organizational error, 26 of the 151 participants who scored above any of the cut-off points (17%) were not invited to the OHS. Consequently, 125 participants (65% of the total group) were invited to the OHS for a consultation.

In the EMH-care group, 61% screened positive on impaired mental health and were invited to follow one or more e-mental health interventions. Those who screened negative on all mental health aspects (39%) were offered to follow *Psyfit,* to retain or enhance their mental fitness. Sixty percent screened positive on impaired work functioning and received the onscreen educational leaflet on how to improve their work functioning. A large overlap existed in those who screened positive on impaired work functioning and those who screened positive on impaired mental health: 139 participants (78%) screened positive on either impaired work functioning or impaired mental health, or both.

Table 2. Baseline characteristics of study participants.

	Occupational physician care (N = 191)		E-mental health care (N = 178)	
	N	(%)	N	(%)
Gender				
Female	156	(82)	147	(83)
Age in years				
18-24	12	(6)	29	(16)
25-34	42	(22)	62	(35)
35-44	44	(23)	24	(14)
45-54	60	(31)	44	(25)
55-64	33	(17)	19	(11)
Occupation				
Nurse	116	(61)	129	(73)
Nurse practitioner	12	(6)	11	(6)
Surgical nurse	12	(6)	0	(0)
Anesthetic nurse	12	(6)	0	(0)
Allied health professionals	29	(15)	31	(17)
Other	10	(5)	7	(4)

After the baseline questionnaire, it was discovered that five men in the OP-care group and twelve men in the EMH-care group who scored above cut-off point on risky drinking behavior accidentally were not classified as such by the technical system. For respectively two and nine of these participants, risky drinking behavior was the only aspect they screened positive on. Risky drinking behavior was not mentioned in their personal results, and – for those in the EMH-care group – they were invited for *Psyfit* instead of *Drinking Less*.

Participants' compliance

The OHS recorded that 51 of the 125 invited participants in the OP-care group (41%) came to the preventive consultation. Most participants (80%) reported not to have gone to the consultation because they regarded it as unnecessary. Of those who went and received an advice from their OP, 80% followed that advice (see Table 3). Reasons reported for not following the advice were that participants regarded it as unnecessary (N = 2) or because of lack of time (N = 1).

Track & trace data showed that 30 of the 178 participants in the EMH-care group (17%) logged on at least once in one or more e-mental health interventions (see Table 4). Six participants logged into two e-mental health interventions. Nine participants (5%) started one or more of the modules of which the e-mental health interventions consisted (six participants started with *Psyfit* and three with *Colour your life*), the other 21 participants merely logged in and did not start any of the interventions' modules.

Four participants in the EMH-care group reported to have followed an e-mental health intervention (see Table 3). Participants chose an e-mental health intervention that fitted their complaints the best. No reasons were reported for not following an e-mental health intervention. Five participants reported to have followed the advice in the onscreen

educational leaflet on how to improve work functioning. Not following the advice because it was regarded as unnecessary, was the most reported reason.

Participants' perspective on the study

In Table 5, the participants' evaluation of the WHS mental module as performed in this study and their perspective on the effectiveness of the interventions is shown. Thirteen percent of participants in the OP-care group and 17% in the EMH-care group would rather receive their personal results differently in the future. Of the participants in the OP-care group who came to the preventive consultation, 79% felt that the personal results reflected their mental health

Table 3. Compliance of participants per study group, based on self-report data.

Occupational physician care	**N**	**(%)**
Participants in OP-care group (N = 191)		
Reported to have received invitation for consultation	123	(100)
Yes	60	(49)
No	63	(51)
Reported to have discussed personal results and/or advice from OP with supervisor	122	(100)
Yes	16	(13)
No	106	(87)
Received invitation from OHS for consultation (N = 125)		
Reported to have gone to consultation	60	(100)
Yes	38	(63)
No	22	(37)
Reasons reported for not going to consultation	20	(100)
Went to consultation (N = 51)		
Reported to have received advice from OP on improving mental health and/or work functioning	35	(100)
Yes	18	(51)
No	17	(49)
Reported to have followed advice on improving mental health and/or work functioning	20	(100)
Yes	16	(80)
No	4	(20)
E-mental health care	**N**	**(%)**
Participants in EMH-care group (N = 178)		
Reported to have discussed personal results and/or advice with supervisor	82	(100)
Yes	9	(11)
No	73	(89)
Received invitation for e-mental health intervention (N = 178)		
Reported to have followed one or more e-mental health interventions	82	(100)
Yes	4	(5)
No	78	(95)
Received onscreen educational leaflet on how to improve work functioning (N = 107)		
Reported to have received onscreen educational leaflet	81	(100)
Yes	23	(28)
No	58	(72)
Reported to have followed advice in leaflet	22	(100)
Yes	5	(23)
No	17	(77)
Explanation on following advice in leaflet	13	(100)

Table 4. Track & trace data on participants' compliance to the intervention.

E-mental health care (N=178)			
	N	**Modules started on**	
		Median	**(range)**
Strong at Work (8 modules)			
Logged on at least once	8	0	(0)
Colour your life (8 modules)			
Logged on at least once	4	2.5	(0-4)
Don't Panic Online (6 modules)			
Logged on at least once	2	0	(0)
Drinking Less (4 modules)			
Logged on at least once	0	0	(0)
Psyfit (6 modules)			
Logged on at least once	22	0	(0-6)
	N	**% followed**	
Psyfit*		**Min.**	**Max.**
Started on module 1 *(Feeling positive)*	4	11	56
Started on module 2 *(Positive relationships)*	1	25	25
Started on module 3 *(Here and now)*	2	32	32
Started on module 4 *(Thinking and feeling)*	2	5	37
Started on module 5 *(Grip on your life)*	3	4	44
Started on module 6 *(Mission and goals)*	2	30	40

* For Psyfit, data are also shown separately per module, as it is the only e-mental health intervention for which modules can be followed separately (e.g. it is not necessary to finish module 1 to start with module 2). The minimum and maximum percentage that was followed of each module, is given for the people who started that specific module.

Table 5. Participants' evaluation and perceived effectiveness of the WHS mental module, self-report data.

Evaluation of the WHS mental module as performed in this study	**Range**	**N**	**(%)**
Occupational physician care			
Receipt of personal results (N = 126)			
Would rather receive personal results differently in the future		16	(13)
Personal results reflected mental health condition (not self-reported) (N = 42)*			
Completely		20	(48)
Partially		13	(31)
No		9	(21)
Personal results reflected work functioning (not self-reported) (N = 46)*			
Completely		26	(57)
Partially		6	(13)
No		14	(30)
Receipt of invitation for preventive consultation (N = 64)			
Would rather receive invitation differently in the future		11	(17)
Preventive consultation			
Felt that they could be open and honest during consultation (N = 38)		37	(97)
Found the advice that was given useful (N = 21)		18	(86)
E-mental health care			
Receipt of personal results (N = 82)			
Would rather receive personal results differently in the future		14	(17)
E-mental health interventions (N = 4)			
Grade given to e-mental health interventions (1, bad – 10, good)	1-8		
Perceived effectiveness of the interventions		**N**	**(%)**
Occupational physician care			
Following OPs advice helped to improve mental health and/or work functioning (N = 15)		9	(60)
E-mental health care			
Following e-mental health intervention(s) helped to improve mental health and/or work functioning (N = 4)		0	(0)
Following advice on how to improve work functioning helped to improve it (N = 5)		3	(60)

* This was not self-reported by the participants, but by the OPs when the participant had come to the consultation.

condition and 70% felt that the personal results reflected their work functioning (partially or completely), according to their OP. Almost all of the participants who came to the consultation (97%) felt they could be open and honest with their OP about their condition. Four participants in the EMH-care group graded the e-mental health interventions (range: 1-8). Nine out of 15 participants in the OP-care group felt that the advice helped to improve their mental health and/or work functioning.

Participants' preferences for future implementation and the WHS mental module is shown in Table 6. In the OP-care group, 79% would or would maybe appreciate to be periodically offered a WHS in the future, versus 74% in the EMH-care group. Most of the participants in both groups (74% versus 67%) would prefer to receive their results on the computer or by e-mail. Preferences regarding how to receive advice to retain mental health and healthy work functioning differed between the study groups. The largest number of participants in the OP-care group (32%) prefer to receive advice through e-mental health interventions, while the largest number of participants in the EMH-care group (37%) prefer to receive advice in a personal consultation with their OP. In both groups, most participants (60% versus 62%) prefer to make an appointment with their OP themselves.

Occupational physicians' protocol adherence

In most consultations, all seven steps of the protocol were followed (Table 7). Expectations about the preventive consultations (step 1) were discussed in each consultation, as well as the possible causes and consequences of impaired mental health and impaired work functioning (step 3). The OPs gave advice in 61% of the 51 consultations. In the remaining 20 consultations, the OPs regarded giving advice to be unnecessary. Ten participants were referred to other healthcare providers, of which four were referred to a company welfare worker and two to a psychologist. Five employees had another follow-up consultation with the OP within six months after baseline.

Occupational physicians' perspective on the WHS mental module

Three of the five participating OPs found that the training prepared them well for the preventive consultations. Three OPs regarded all steps of the protocol for the consultations as useful and performable. All in all, the protocol was regarded to be of use and to fit their method of working.

Discussing the personal results seemed to be sensitive, due to participants' lack of identification with their personal results. The OPs felt it was important to clarify to the participants that their symptoms might be very light, thus might not be experienced as an impairment. However, the OPs also doubted the quality of the screening instruments and cut-off points that were used in the study, particularly the one used for risky drinking behavior.

Table 6. Participants' preferences regarding future implementation of a WHS mental module, self-report data.

	Occupational physician care		E-mental health care	
	N	(%)	N	(%)
Would appreciate periodical offer of WHS	112	(100)	69	(100)
Yes	41	(37)	23	(33)
Maybe	47	(42)	28	(41)
No	24	(21)	18	(26)
Would participate in a WHS aimed at mental health and work functioning	107	(100)	68	(100)
Yes	39	(36)	23	(34)
Maybe	51	(48)	33	(49)
No	17	(16)	12	(18)
Preference to receive personal results	107	(100)	64	(100)
Immediately by the computer and/or by e-mail	79	(74)	43	(67)
In a personal consultation with OP	17	(16)	17	(27)
Otherwise	11	(10)	4	(6)
Preference to receive tailored advice	104	(100)	62	(100)
Personal consultation with OP	29	(28)	23	(37)
E-mental health interventions	33	(32)	16	(26)
Personal consultation with OP with opportunity to follow e-mental health interventions	23	(22)	14	(23)
Through other in-house healthcare providers	5	(5)	1	(2)
Through other external healthcare providers	2	(2)	1	(2)
Preference for a consultation with OP	52	(100)	37	(100)
Make an appointment	31	(60)	23	(62)
Receive an invitation for an appointment which is made	21	(40)	14	(38)

More than half of the participants did not come to the appointment, making it difficult for the OPs to schedule their time efficiently. A solution would be to let employees make their own appointment.

The OPs perceived the supervisors' and managers' knowledge and acceptance of the study to be marginal, and found that the organization did not seem ready for implementing a WHS mental module. According to the OPs, the resistance of supervisors and managers stemmed from fear to be complained about during the consultation with the OP.

The OPs regarded the preventive consultations to be meaningful. For roughly 50 percent of the consultations, the OPs could mean something to the employees immediately. However, for the other half of consultations it was also useful, as an introduction to occupational healthcare and to improve the image of the OHS.

The OPs were asked about their perspective and preferences regarding future implementation of a WHS mental module. Most OPs regarded future implementation as meaningful if the WHS mental module would be adapted. Most regarded the chance to be considerable that implementation would be successful. The in-house communication could be improved, by repeatedly informing all layers of the organization and assessing whether the message has come across. The OPs thought that they should play an active role in this

Table 7. Adherence of OPs to protocol during preventive consultations (N = 51).

Elements of protocol adherence	Median	(Range)
Adherence to protocol steps		
Number of protocol steps followed	7	(4-7)
	N	**(%)**
Adherence to protocol, per step		
Step 1: Expectations – discussed	51	(100)
Step 2: Screening results – discussed	50	(98)
Step 3: Causes and consequences – discussed	51	(100)
Step 4: Problem diagnosis & rationale – discussed	46	(90)
Step 5: Advice – discussed	45	(88)
Step 6: Follow-up – discussed	35	(69)
Step 7: Summary – discussed	50	(98)
Advice given to employee		
Yes, advice given	31	(61)
Yes, aimed at both the person/the complaints and work functioning	*15*	*(29)*
Yes, aimed at work functioning	*11*	*(22)*
Yes, aimed at the person/the complaints	*5*	*(10)*
No, no advice given	20	(39)
Advice was not needed	*20*	*(39)*
Did not get round to this step	*0*	*(0)*
Referral of the employee		
Employee was referred to other healthcare provider	10	(20)
Follow-up		
Follow-up appointment not planned	43	(86)
If needed, the employee can make another appointment	4	(8)
Follow-up consultation planned within next (6 weeks; 4 weeks)	2	(4)
Contact by telephone/e-mail within next 4 weeks	1	(2)

communication and that they could be responsible for personally informing the medical divisions and wards. Another point to consider according to the OPs, was the phrasing of personal results which could be improved in the future to facilitate employees' interpretation.

DISCUSSION

Response of the OP-care group and the EMH-care group was 32%. Forty-one percent of the invited participants in the OP-care group went to the preventive consultation and 80% followed the OP's advice if this was given. Seventeen percent in the EMH-care group logged into an e-mental health intervention and 5% started it, at least to some extent. The advice on improving work functioning, given in an onscreen educational leaflet, was followed by 23% of the participants who received it. Thirteen percent of participants in the OP-care group and 17% in the EMH-care group would rather receive their personal results differently in the future. The OPs rated that 70-80% of the participants who came to the consultation, felt that the personal results reflected their work functioning and mental health. Moreover, of the participants who came to the consultation, almost everyone felt they could be open and honest with their OP. The preventive consultation was perceived effective by nine out of 15 participants. In both

groups, most participants would appreciate to be offered a WHS in the future. The adherence of OPs to the consultation protocol was high, as in most cases all steps were followed. The OPs were satisfied with the consultation protocol and training, and felt that the preventive consultations had been meaningful. They also considered it to be meaningful to implement a WHS mental module in the future.

For each participant who came to the preventive consultation, OPs registered whether the participants felt that their personal results reflected their mental health condition and work functioning. This was used as one of the process indicators of participants' evaluation of the WHS mental module. It would have been more accurate to assess this indicator at the employee-level rather than the OP-level for two reasons: the information would have been received of both groups and not only of the OP-care group, and it would have given a more accurate evaluation, as it would not have been given through another person.

Recruitment for the study was time-consuming. However, it was found useful, as a surge in participation was observed after most of the recruitment activities. The response rate in this study is similar to other studies in which employees were screened on mental health complaints[27-29] or on general health risks.[30]

Some of the reasons that were given for not participating in this study were: feeling overloaded with research, because the employees of the academic medical center are asked to take part in many other research projects; fear of privacy, although anonymity of their enrollment was stressed; and the extensive questionnaire (it took about 30 minutes to fill out). Furthermore, in general, personal benefit is important in the decision to participate in research.[31] The study might have attracted employees who felt that reflection on work functioning and mental health and a possible intervention would be useful or necessary to them. This idea is supported by the high percentages in both groups who screened positive on impaired work functioning and/or impaired mental health. This might have influenced the effect of the interventions on the outcome results.

An important aspect to discuss here, is the large difference between the two study groups in compliance of the participants to the intervention. In the OP-care group, compliance was fairly high. It was not expected that everyone would come to the consultation with their OP when invited, as the offer of the consultation had a preventive nature. However, compliance in the EMH-care group was low. Self-help e-mental health interventions often have shown to have low compliance rates.[32] In a study with participants who wanted to work on their depressive complaints,[14] it was found that of the people who were randomized to follow the e-mental health intervention *Colour your life,* 36% followed it to an adequate extent (at least five sessions). In a further study,[33] several barriers were identified which participants of *Colour your life* experienced, such as experiencing a lack of identification with and applicability of the

e-mental health intervention, and inadequate computer and Internet skills, equipment, or location. It is possible that these aspects played a role in the low compliance of the nurses and allied health professionals in this study. We observed that some of the participants experienced problems with logging into the interventions, due to inadequate computer skills, but also due to technical problems. A further barrier might have been that nurses and allied health professionals are very people-oriented employees, possibly making e-mental health interventions too impersonal for them.

Lastly, the present study regarded a preventive setting in which the e-mental health interventions were offered without the participants specifically asking for any intervention. To our knowledge, using e-mental health interventions in a preventive WHS setting has not been researched before. It should be investigated how the e-mental health interventions can be improved to make them more appealing for nurses and allied health professionals and to increase compliance. Furthermore, analyses of the (cost-)effectiveness of the Mental Vitality @ Work program will determine whether the interventions employed in this study were sufficient to improve help-seeking behavior, work functioning, mental health and well-being.

A possible improvement of the WHS mental module would be to combine the two strategies that were compared in this study. The employee could be screened online, and given feedback on their personal results onscreen or through e-mail as found preferred by the participants. They could then either choose to see their OP in a preventive consultation, or to start an e-mental health intervention. For a preventive consultation, the participants in the present study prefer to make an appointment with their OP themselves, which was also suggested by the OPs as an adaptation of the WHS mental module. The combination of care by the OP and e-mental health care would fit the participants' preferences found in this study. In the preventive consultation, the OPs could also offer an e-mental health intervention (possibly under the guidance of the OP), if this is deemed appropriate and if it suits the employee's preference. This would also make the WHS mental module more personal and possibly more suited to nurses and allied health professionals.

Conclusions

It is possible to include a WHS mental module for nurses and allied health professionals in the OHS' activities. The WHS mental module was well-received by the nurses and allied health professionals who participated in the study. It should be investigated how to improve the response and the compliance to the intervention.

REFERENCES

1 Wieclaw J, Agerbo E, Mortensen PB et al. Risk of affective and stress related disorders among employees in human service professions. *Occup Environ Med* 2006;63:314-9.
2 Gärtner FR, Nieuwenhuijsen K, van Dijk FJ et al. The impact of common mental disorders on the work functioning of nurses and allied health professionals: A systematic review. *Int J Nurs Stud* 2010;47:1047-61.
3 Nieuwenhuijsen K, Verbeek JH, de Boer AG et al. Validation of performance indicators for rehabilitation of workers with mental health problems. *Med Care* 2005;43:1034-42.
4 van der Klink JJ, Blonk RW, Schene AH et al. Reducing long term sickness absence by an activating intervention in adjustment disorders: a cluster randomised controlled design. *Occup Environ Med* 2003;60:429-37.
5 Koh D, Aw TC. Surveillance in occupational health. *Occup Environ Med* 2003;60:705-10.
6 Plat MJ, Frings-Dresen MH, Sluiter JK. A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *Int Arch Occup Environ Health* 2011;84:839-57.
7 Linnan L, Steckler A. Process evaluation for public health interventions and research: an overview. Process Evaluation for Public Health Interventions and Research. San Francisco: Jossey-Bass, 2002: 1-23.
8 Bouffard JA, Taxman FS, Silverman R. Improving process evaluations of correctional programs by using a comprehensive evaluation methodology. *Evaluation and Program Planning* 2003;26:149-61.
9 Nielsen K, Fredslund H, Christensen KB et al. Success or failure? Interpreting and understanding the impact of interventions in four similar worksites. *Work & Stress* 2006;20:272-87.
10 Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. *Qual Saf Health Care* 2003;12:40-6.
11 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.
12 van der Klink JJ. Guideline: Management of mental health problems of workers by occupational physicians *[in Dutch: Richtlijn: Handelen van de bedrijfsarts bij werkenden met psychisch problemen]*. Utrecht: Netherlands Society of Occupational Medicine *[in Dutch: NVAB]*, 2007.
13 Bolier L. Haverman M, Bohlmeijer E, Psyfit online - Randomized effect study on an online mental fitness self-help program - Internal report [*in Dutch: Psyfit online - Gerandomiseerde effectstudie naar een e-mental fitness zelfhulpprogramma - Intern rapport*] Utrecht: Trimbos-instituut, 2010.
14 de Graaf LE, Gerhards SA, Arntz A et al. Clinical effectiveness of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry* 2009;195:73-80.
15 de Graaf LE, Huibers MJ, Riper H et al. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. *J Affect Disord* 2009;116:227-31.
16 Spek V, Nyklicek I, Smits N et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. *Psychol Med* 2007;37:1797-806.
17 Spek V, Cuijpers P, Nyklicek I et al. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007;37:319-28.
18 Spek V, Cuijpers P, Nyklicek I et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. *Psychol Med* 2008;38:635-9.
19 Warmerdam L, van Straten A, Twisk J et al. Internet-based treatment for adults with depressive symptoms: randomized controlled trial. *J Med Internet Res* 2008;10:e44.
20 Meulenbeek P, Willemse G, Smit F et al. Early intervention in panic: randomized controlled trial and cost-effectiveness analysis. *Trials* 2008;9:67.
21 Meulenbeek P, Willemse G, Smit F et al. Effects and feasibility of a preventive intervention in sub-threshold and mild panic disorder: Results of a pilot study. *BMC Res Notes* 2009;2:4.
22 Meulenbeek P, Willemse G, Smit F et al. Early intervention in panic: pragmatic randomised controlled trial. *Br J Psychiatry* 2010;196:326-31.
23 Smit F, Willemse G, Meulenbeek P et al. Preventing panic disorder: cost-effectiveness analysis alongside a pragmatic randomised trial. *Cost Eff Resour Alloc* 2009;7:8.
24 van Ballegooijen W., Riper H, van Straten A. et al. The effects of an Internet based self-help course for reducing panic symptoms - Don't Panic Online: study protocol for a randomised controlled trial. *Trials* 2011;12:75.

25 Kramer J, Riper H, Lemmers L et al. Television-supported self-help for problem drinkers: a randomized pragmatic trial. *Addict Behav* 2009;34:451-7.
26 Riper H, Kramer J, Smit F et al. Web-based self-help for problem drinkers: a pragmatic randomized trial. *Addiction* 2008;103:218-27.
27 Harvey SB, Glozier N, Henderson M et al. Depression and work performance: an ecological study using web-based screening. *Occup Med (Lond)* 2011;61:209-11.
28 Lexis MAS, Jansen NWH, Huibers MJH et al. Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial. *Occup Environ Med* 2011;68:400-7.
29 Wang PS, Simon GE, Avorn J et al. Telephone Screening, Outreach, and Care Management for Depressed Workers and Impact on Clinical and Work Productivity Outcomes. *JAMA* 2007;298:1401-11.
30 Colkesen EB, Kraaijenhagen RA, Frings-Dresen MHW et al. Participation in a workplace web-based health risk assessment program. *Occup Med (Lond)* 2011;61:586-9.
31 Cook AF, Hoas H. Trading Places: What the Research Participant Can Tell the Investigator about Informed Consent. *J Clinic Res Bioeth* 2011;2:121-7.
32 Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7:e11.
33 Gerhards SAH, Abma TA, Arntz A et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. *J Affect Disord* 2011;129:117-25.

9

General discussion

The main objectives of this thesis were to develop and evaluate a job-specific instrument to assess work functioning impairments due to common mental disorders (CMDs) in nurses and allied health professionals (research questions 1 to 3) and to develop and evaluate a workers' health surveillance (WHS) mental module for nurses and allied health professionals (research questions 4 and 5). In this chapter, the main findings are summarized, the methodological strengths and limitations are discussed, and reflections on the main findings are presented. At the end of this chapter, recommendations for future research and practice are provided.

MAIN FINDINGS

1) Aspects of work functioning that can be impaired due to common mental disorders in nurses and allied health professionals

The systematic review and five focus groups with nurses and allied health professionals and occupational health experts yielded 13 aspects of impaired work functioning, which are regarded assessable by a self-report questionnaire. Overall, these aspects concern impairments in task-related functioning aspects, causing incidents or near-misses, intrapersonal, and interpersonal aspects of work. In the systematic review, half of the aspects found were supported only by narrative evidence, as the association between these aspects and CMDs was not corroborated by quantitative data. The systematic review revealed no study that specifically focused on allied health professionals and work functioning impairments due to CMDs. The findings yielded from the literature review and the focus groups overlapped to a large extent. However, the focus group data provided more details and examples of behavior for all inventoried aspects.

2) The Nurses Work Functioning Questionnaire: content validity, factorial structure, and internal consistency

A questionnaire with seven subscales and a total of 50 items called the Nurses' Work Functioning Questionnaire (NWFQ) was established. The seven subscales are as follows: 1) *Cognitive aspects of task execution and general incidents*, 2) *Impaired decision making*, 3) *Causing incidents at work* (not suitable for allied health professionals), 4) *Avoidance behavior*, 5) *Conflicts and annoyances with colleagues*, 6) *Impaired contact with patients and their family*, and 7) *Lack of energy and motivation*. The factorial structure was developed by explorative factor analysis and corroborated in confirmatory factor analysis. Regarding content validity, all aspects were evaluated as relevant or highly relevant, and the representativeness of the item pool was assessed as highly comprehensive by six experts. Among the seven definitive subscales, internal consistency was good in four subscales (Cronbach's alphas ranged from 0.81 to 0.94) and acceptable in three subscales (Cronbach's alphas ranged from 0.70 to 0.78).

3) Clinimetric quality of the NWFQ in terms of reproducibility, construct validity, and the interpretability of change

To determine the reproducibility of the NWFQ, the reliability and the level of agreement were evaluated in 112 nurses and allied health professionals. All subscales showed good reliability with intraclass correlation coefficients (ICC) above 0.70, except for subscale 2), *Impaired decision making*, for which the ICC was too low (ICC = 0.16). For six subscales, the standard error of measurement (SEM) ranged from 3 to 6, which is considered low for a scale range from 0 to 100. However, the SEM for subscale 2 was rather high (SEM = 17). All subscales reached the hypothesized values for good convergent, divergent, and discriminative validity except for subscale 2, which showed good divergent and discriminative validity but did not meet the criterion for good convergent validity. It is concluded that the NWFQ demonstrates good reproducibility and construct validity for six of the seven subscales, which are regarded suitable for future use. Subscale 2), *Impaired decision making*, needs improvement before further use. Regarding the interpretability of change, it is concluded that three subscales exhibit good interpretability of change. Further analyses of the interpretability of change and of the minimal detectable change values are recommended, as the study sample consisted of a low number of subjects who improved in their work functioning during the three-month period. Provisionally, the calculated MIC values can cautiously be used.

4) The effectiveness of a workers' health surveillance mental module for nurses and allied health professionals

In total, 191 nurses and allied health professionals took part in a WHS mental module, which included a screening for work functioning and mental health problems. In the control group, 188 workers participated at baseline. Workers who were screened as positive in the intervention group were invited for a consultation with the occupational physician. The effectiveness of the WHS mental module was compared to a control group who were not screened or invited to consult with the occupational physician in a randomized controlled trial. In the group of workers that received the WHS mental module (WHS OP-care group), 151 (79%) of the workers were classified as having work functioning impairments and/or mental health complaints compared to 161 (86%) in the control group.

A statistically significant interaction effect was found for study-group*time on help-seeking behavior (p = 0.02), thus the time course of help-seeking behavior differed significantly for the two study groups. In the WHS OP-care group help-seeking remained stable between baseline and three months follow-up, while it decreased for the control group. The difference between the study groups was 13% at three months of follow-up, which was marginal significant (p = 0.09). Towards six months follow-up, help-seeking behavior decreased more steeply in the WHS OP-care group than it did in the control group, resulting in even slightly less help-seeking behavior. However, the difference was not statistically significant (p = 0.14).

Workers who received the WHS mental module showed statistically less work functioning impairments at three and six months of follow-up when compared to the control group ($p = 0.04$). At three months of follow-up, 45% of the workers who received the intervention showed at least minimal important improvements in work functioning; this was 30% in the control group ($p = 0.03$). At the six-month follow-up, 41% of workers in the WHS OP-care group exhibited improvements representing a change that was minimal or larger; the magnitude of this change in the control group was 28%. This difference was marginal significant ($p = 0.05$).

No significant difference in mental health complaints was found between workers who received the WHS mental module and a control group, except for risky drinking behavior. An interaction effect of study-group*time was found for risky drinking behavior ($p < 0.01$). The difference between the two study groups in risky drinking behavior was not statistically significant at three months ($p = 0.08$) or six months of follow-up ($p = 0.36$), but the timeline of drinking behavior differed. At three months of follow-up, the WHS OP-care group showed a slight rise in drinking behavior compared to the control group; however, at the six-month follow-up, the level had dropped below that observed in the control group.

5) Evaluation of the workers' health surveillance mental module in terms of response, compliance, adherence, and perspectives on the WHS mental module

The process evaluation of the WHS mental module revealed that the response rate was 32%. With respect to compliance, 51 (34%) of the 151 workers in the WHS OP-group, who were screened as positive, followed up upon the invitation for the preventive consultation. Due to system errors, only 125 of the 151 workers were invited for a consultation, thus 41% of the invited workers followed up upon the invitation. 80% of these workers reported to have followed the occupational physician's advice if any had been given. It was found that 13% of the participants in the WHS OP-care group would rather receive their personal results differently in the future. The occupational physicians rated that 70% of the participants who went to the consultation felt that the personal results (partly) reflected their level of work functioning; this value was 60% for the mental health screening results. Moreover, almost all workers who visited the occupational physician for the preventive consultation felt that they could be open and honest with their occupational physician. The preventive consultation was perceived as effective by nine out of 15 participants. Most participants would appreciate to be offered a WHS in the future.

The adherence of the occupational physicians to the consultation protocol was high, as all steps were followed in most cases. The occupational physicians were satisfied with the consultation protocol and training. They felt that the preventive consultations had been meaningful and that implementing a WHS mental module in the future would also be meaningful. The occupational physicians expressed some reluctance and helplessness regarding providing advice and initiating further care in cases when the worker did not recognize the results of the mental health screening. Suggestions were given for improvement

of the present WHS mental module, e.g., regarding the in-house communication strategy and the role of e-mental health interventions.

METHODOLOGICAL CONSIDERATIONS

The development of the NWFQ

One of the strengths of this thesis is the method applied for the NWFQ development. The development followed a clear step-by-step procedure which was planned in advance. In the development process, findings from the literature as well as qualitative data conveying the knowledge and experiences of employees and experts were used as input. Furthermore, in the quality assessment of possible items and subscales, both expert opinions and statistical analyses were used. In conclusion, the procedure employed exemplifies the requirements for the development of a scientific questionnaire that is relevant for practice according to the psychometric[1] and clinimetric literature.[2] Due to the transparency of this comprehensive strategy, it might function as an example of good practice for the development of scientific questionnaires in the field of occupational health.

The underlying construct of the NWFQ is *impaired* work functioning due to CMDs in nurses and allied health professionals. Therefore, in the formulation of items, the focus was placed strongly on aspects of work functioning that are signs of impairment. As a result, items were often formulated *negatively*, in the sense that they present examples of undesirable behavior. This structure is similar to health complaint scales, where items mostly present symptoms of health complaints, and also similar to some existing work functioning questionnaires, such as the Work Limitation Questionnaire (WLQ).[3] On these types of items, the majority of workers will report no impairments because they are healthy and well-functioning and thus, will have a score of zero or slightly above. Two consequences of this characteristic of the NWFQ scores are considered. The first consequence is a non-normal distribution of NWFQ scores when applying the questionnaire in a nursing population comprising mainly healthy and well-functioning workers. The second consequence is that no gradation for *good* work functioning can be assessed using the NWFQ nor can any improvements in work functioning over time among workers who already function well. Therefore, the NWFQ is restricted in its usefulness, e.g., for the human resource management setting. When applying the NWFQ for evaluative purposes in research, these consequences must be considered; however, for detection purposes, as in the context of WHS, they are not expected to have a negative impact.

Case classification by the NWFQ

Due to the absence of a gold standard, the cut-off values for the WHS screening were based on statistical norms, according to the principle that sum scores can lead to three categories:

green, orange, and red. Therefore, two cut-off values were set per subscale based on the percentiles in a population with mild mental health complaints. For the total NWFQ, a red score on one subscale or three or more orange scores led to case identification for impaired work functioning. The method applied is rather arbitrary, but it is common for questionnaires measuring fairly new constructs, for which a gold standard or any other reference standard is nonexistent. One example is the Maslach Burnout Inventory, a widely used self-report questionnaire for the measurement and diagnosis of burnout. For the Dutch version, the Utrecht Burnout Scale, cut-off values were initially based on percentiles.[4] Additional studies for clinically derived cut-off values were recommended by the same authors, which were published in 2001,[5] six years after the questionnaire was first published.

Likewise, further validation of the cut-off values of the NWFQ is necessary. However, the absence of a gold or a reference standard makes a number of regular procedures impossible. According to Rutjes and colleagues,[6] other types of evaluation must be considered in cases where no single reference standard or set of tests for classification exists. One solution might be the use of research methods that were originally developed for other purposes. A method analogous to the anchor-based method for minimal important change value identification might be interesting in this context. Anchor scores could be used to assess whether the work functioning of a workers is perceived as sufficient or impaired. These anchor scores could then be used as a reference standard to evaluate the existing cut-off points using the ROC-curve method or to identify new cut-off points. However, anchor formulation and selection of the right group to fill out the anchor questions represent substantial challenges. Another conceivable option with which to evaluate or identify new cut-off values would involve use of the Delphi method.[7]

For future use of the NWFQ, I assume that the previously identified cut-off values might be sufficient for research purposes, e.g., the dichotomization of data. For screening purposes, they should, however, only be used in combination with a second-stage screening, such as a consultation with an occupational physician or another occupational health professional, e.g., an occupational health nurse, psychologist, social worker, or occupational therapist, depending on the occupational health system of the specific country or company.

Screening strategy for the workers' health surveillance mental module

The screening strategy applied in the WHS mental module led to a high number of positives among the workers screened: 79% in the WHS OP-care group and 86% in the control group. Methodological and theoretical choices for the screening may have influenced this high percentage. First, the applied cut-off values, which were all validated except for the NWFQ (as discussed above), influenced these percentages. For the WHS mental module, cut-off values were selected with a choice for optimal sensitivity but at the cost of specificity. With an eye toward prevention, it was desired not to miss the chance to offer workers the opportunity for a preventive consultation with the occupational physician. Second, the screening results were

supposed to be as detailed as possible to allow for optimal feedback and, therefore, purposeful interventions by the occupational physician. Therefore, it was our intention to separately screen for common mental health problems which can be related to the work of healthcare workers. Also, the various aspects of work functioning impairment were screened for separately. This resulted in the use of six screeners for mental health complaints, and seven aspects of work functioning impairment that were investigated separately. The use of thirteen different screening instruments must have influenced the overall percentage of positive screened workers.

The question rises as to whether the right group of workers was identified. Not all workers who were screened as positive can be seen as workers with CMDs or serious work functioning problems. However, all of these workers had reported symptoms of either work functioning impairments or a mental health complaint, which were at least classifiable as mild to moderate severe symptoms. Therefore, the individuals labeled as positive represent a group of workers in which the improvement of work functioning and/or mental health should be possible.

Nevertheless, possible harms of the applied screening strategy should be considered. The screening method applied was regarded as safe for two reasons. First, the personal feedback on screening results that was provided online was formulated mildly without speaking of *health problems* or *diagnosis* to prevent any wrong interpretations or premature conclusions. Second, the consultation with the occupational physician was regarded as a second screening-stage; no treatment had been initiated before the visit to the occupational physician (or any other caregiver). Only a few notifications were received from supervisors and workers who were upset by the personal results. This was mainly related to the risky drinking behavior scale. However, as part of the process evaluation, the occupational physicians reported that for 70% of the workers who came to the preventive consultation, screening results (partially) reflected their perceived work functioning, and in 60% of the workers, screening results (partially) reflected their mental health state (Ketelaar et al., submitted). This should be regarded as satisfactory, as difficulties with recognizing mental health complaints discovered through a screening is not uncommon.[8] Furthermore, among the workers who followed the invitation for the occupational physician consultation, 85% were content with the advice they received on improving work functioning and/or mental health.

In future improvements of the WHS mental module, the choice for cut-off values and for the feedback strategy should be reevaluated to further reduce any possible harm. One option is to apply cut-off values with high sensitivity but low specificity; however, in this case, personal screening results should be presented in a safer and more personal environment, e.g., by the occupational physician or an occupational health nurse. Another option is to present the results of the screening online, as was done in this study. In such a case, it would be necessary to apply less sensitive cut-off values to even more purposively avoid the negative effects of feedback in false-positive cases. This second strategy would be preferable in light of

the process evaluation, which showed that employees prefer online feedback to receiving personal feedback from an occupational physician (Ketelaar et al., submitted).

INTERPRETATION OF AND REFLECTION ON THE MAIN FINDINGS

The additional value of the NWFQ

In the introduction of this thesis, the need for more insight into job-specific aspects of work functioning impairments due to CMDs was formulated, as well as the necessity for a work functioning questionnaire that operationalizes impaired work functioning due to CMDs in nurses and allied health professionals. These aims were achieved successfully in the first four studies of this thesis. The conceptual model of work functioning described in the General introduction (Chapter 1), was initially designed as a tool to support the development process of the questionnaire. As could be expected, the subscales identified for the NWFQ represent all parts of work functioning as defined in the conceptual model, which were task and contextual elements, as well as the four following dimensions: process, quality outcomes, quantity outcomes, and effort. It has to be noted, that the conceptual model was designed with the development of a questionnaire for impaired work functioning of hospital workers in mind. Thus, before applying the model for other aims, it should be reevaluated and possibly adapted for the specific purpose.

It is concluded that the NWFQ is of additional value compared to existing generic work functioning questionnaires. The first feature in which the NWFQ differs is that it examines aspects of work functioning that are not, or are to a lesser extent, included in generic work functioning instruments. One of these aspects is causing incidents at work, an important risk of working in healthcare. The other aspect is interpersonal behavior, which in the NWFQ is separately measured for contact with colleagues and for communication with patients and their family. The second feature that exemplifies the added value of the job-specific questionnaire is the formulation of its items, which mainly presents concrete examples of behaviors and tasks. I assume that therefore less interpretation of the items investigated is necessary, and reflection on one's own behavior is facilitated; both benefits promote the accuracy of the self-report measure.

The scope of application of the NWFQ

The NWFQ is innovative because it measures work functioning impairments that are both disease- and job-specific. This specific approach might influence the scope of application of the questionnaire. In the following text, I will reflect on the applicability of the NWFQ within the context of WHS, within occupational health practice for purposes other than WHS, and for research purposes.

Use of the NWFQ in workers' health surveillance

The NWFQ encompasses various signs for impaired work functioning, which makes the questionnaire advantageous for both screening and the intervention in a WHS. For screening purposes, it allows for the classification of workers who demonstrated impaired work functioning during a certain period compared to other periods in their careers. However, in the future, cut-off points should be further validated to reduce the risk of misclassification.

With regard to the intervention element in workers' health surveillance, NWFQ scores provide a good starting point for the occupational physician to select interventions that can be discussed with the worker. The multidimensionality and the concrete examples of work behaviors and tasks that are reflected in the items provide detailed insight into the exact aspects of work that are presumably impaired. This insight offers input on work aspects that might be necessary to adapt in favor of the health and safety of the workers and their patients. To reduce short-term risks for health and safety, temporary changes in the work content, work organization (e.g., working with or without a colleague on specific tasks), type of shifts, or number of work hours might be appropriate. Over the long run, one should focus on the adaptation of the workers' perception, work methods, and work conditions. As a positive effect was found for the WHS mental module on work functioning impairments, it can be assumed that the NWFQ scores as part of the screening results do indeed contribute to those adaptations.

Use of the NWFQ in occupational health care

Use of the NWFQ in occupational health practice, outside of the WHS setting, remains to be investigated. It would be interesting to further investigate the value of the NWFQ for evaluative purposes in counseling by occupational physicians, e.g., of nurses who return to work after a period of sickness absence. Workers who return to work after sickness absence spells due to CMDs are often not fully recovered.[9 10] Therefore, impairments in work functioning are still an issue after work resumption.[11] For this group of workers, detailed insight into work functioning impairments may be useful to prevent risks for themselves and for the patient. The results of the NWFQ, e.g., filled out online on a regular basis, might be of value for the identification of difficulties in the work exertion and for insights into which aspects of work might need further attention and guidance. To interpret changes in work functioning over time, MIC values of the NWFQ may be an indication of important improvements.

Use of the NWFQ in research

With a shift in focus from absenteeism to work functioning (presenteeism) in the occupational health research during the last decade, work functioning is used more often as an outcome measure in observational and interventional studies. The NWFQ provides a suitable measure for these aims in studies on nurses and allied health professionals. However, the application range for research purposes might be reduced because the NWFQ operationalizes *impaired*

work functioning, and consequently, it does not assess differences in the functioning of workers who have no or only small impairments. Application of the questionnaire in a population with mainly healthy and well-functioning workers therefore is not sensible, except for screening purposes. Furthermore, as mentioned above, the NWFQ scores do not follow a normal distribution in the general population of nurses and allied health professionals. Therefore, NWFQ scores need to be adapted by log-transformation, which, however, reduces the ease of interpretability of results. Alternately, non-parametric analyses could be performed, which, however, result in a decrease in power.

Not all scientific studies have the scope to evaluate multiple aspects of impaired work functioning separately. Using the total score of the NWFQ might be a good alternative when insight into the separate subscales is not possible or relevant. In post-hoc analyses, we calculated the intraclass correlation coefficient of the total NWFQ score (without the impaired decision making subscale) and found it to be above 0.90, a value that met the criterion for good reliability. Thus, the reliability of the NWFQ total score is regarded as good.

The NWFQ: a template for further work functioning questionnaires?

The studies on the NWFQ might function as a template for the development of work functioning instruments for other occupations. For occupations with a high prevalence of CMDs, in which work functioning impairments are associated with risks for the health and safety of the worker and the safety and well-being of others, comparable questionnaires might be relevant. This is the case for healthcare workers other than nurses and allied health professionals, such as doctors or care-givers outside the hospital setting, but also for human service workers such as teachers, social workers, and emergency responders, e.g., firefighters, ambulance workers, and police staff.[12]

For other occupations in the healthcare sector, it might be sufficient to adapt the NWFQ by making small adjustments to certain items to better reflect the daily work of the specific occupational group. For occupations outside of healthcare service, the methods section of Chapter 3 of this thesis provides an example of good practice for the scientific development of a job-specific and health-related work functioning questionnaire. However, the applied approach is comprehensive and time consuming and may therefore not always be feasible. The systematic literature review yielded no themes of impaired work functioning due to CMDs that were not also mentioned during the focus groups. Additionally, the focus group interviews gave more detailed insight into the exact tasks and behaviors that are impaired. Therefore, for a less time-consuming development process, it is advisable to combine a quick literature search with thorough qualitative data gathering (through the use of semi-structured in-depth interviews or focus group interviews).

The subscales that are regarded as adaptable to most jobs in the human service sector concern the following aspects: cognitive aspects of work; impaired decision making; avoidance behavior; contact and communication with co-workers, as well as with other

persons (e.g., pupils and parents for teachers and the public for police officers); and lack of energy and motivation. The subscale *Causing incidents at work* is regarded as crucial for most occupations that have responsibilities for the safety and well-being of others; however, it was difficult to apply this subscale to the evaluation of occupational groups for which reporting incidents is less common (allied health professionals). One alternative would be to omit the words "incidents" and "near-misses" and instead describe concrete adverse events specific to an occupational group. In such a questionnaire, e.g., for police officers, this could mean the failure to act immediately upon a situation or to judge the seriousness of a situation incorrectly.

The effectiveness of the workers' health surveillance mental module

The WHS mental module developed and evaluated as part of this thesis aimed to stimulate help-seeking behavior in workers with work functioning impairments and/or mental health complaints. An increase in the number of workers who have visited a caregiver was observed. Although, help seeking increased only by following the invitation for the preventive consultation with the occupational physician, not for other type of caregivers. An underlying presumption was that providing feedback on screening results would enhance help-seeking behavior, due to an increase of recognition of mental health complaints and impairments in work functioning. However, effects on the recognition of complaints were not measured directly. Knowledge of the possible effects on recognition would be useful, as it could provide further insight into the stimulating effect on self-reflection by the feedback of screening results or by the preventive consultation with the occupational physician.

Furthermore, for this WHS mental module, the underlying assumption was that in consultation with the occupational physician (or any other caregiver) self-evidently interventions would be initiated. This would in turn lead to an improvement in work functioning and a reduction of mental health complaints. However, effects were only found on work functioning and on risky drinking behavior.

A decrease in work functioning impairments was observed among the group of workers to which the WHS mental module was offered. This effect on work functioning is promising, especially when considering that only 41% of the 125 positively screened workers who received an invitation to consult with their occupational physician accepted the invitation. Although the effect of the WHS mental module on the improvement of work functioning is promising, still, little insight has been gained into what establishes this effect. It might be based on the personal feedback, the consultation with the occupational physician, or its interplay, as well as on work adaptations or changes in behavior initiated by the worker him- or herself. Furthermore, as a total score for work functioning was used as outcome measure, no effects for the various different aspects of work functioning have been investigated. Future

research should focus on understanding the sources of the effect on work functioning. With this insight, the protocol for the preventive consultation could be improved.

No effects of the intervention were found on the measured mental health complaints, other than risky drinking behavior. Risky drinking behavior as measured in the effect study of the WHS mental module might be regarded as a behavioral outcome rather than a mental health complaint. Therefore, it might be concluded that the WHS mental module has an effect on behavioral changes, i.e., work functioning and risky drinking behavior, whereas not on the mental health state. The process evaluation of the WHS mental module showed that the participating occupational physicians doubted the quality of the screeners and cut-off points used in the study to some extent, particularly those regarding risky drinking behavior. However, the screening results regarding risky drinking behavior were extensively discussed in most preventive consultations, although without providing advice for the regulation of alcohol consumption. Therefore, it might be assumed that the attention given to risky drinking, even without providing advice, had an effect on this behavior.

One possible explanation for the absence of a reduction in mental health complaints might be a lack of readiness to offer preventive treatment for mental health complaints on the part of the occupational physicians, when confronted with workers who did not perceive a mental health problem. A focus group was held with the participating occupational physicians as part of the process evaluation (Chapter 8). The occupational physicians expressed reluctance and helplessness regarding providing advice and initiating further care in cases when the worker did not recognize the results of the mental health screening. The process evaluation also showed that only a limited number of workers were scheduled for follow-up appointments with the occupational physician. Reluctance on the part of primary care providers with regard to providing brief interventions for preventive purposes is not uncommon, as a study on alcohol use screening and a subsequent brief intervention by general practitioners showed.[13] Issues such as the caregivers' own perception of screening and the stigma related to the screened complaints are some of the factors underlying this reluctance.[13] Other reasons for reluctance reported included concerns about the relationship with the client and the doctors' understanding of their work as being primarily treatment rather than prevention.

To stimulate more proactive preventive guidance, the training for the WHS mental module administered to occupational physicians should more strongly emphasize that, for preventive purposes, mild health complaints do not necessarily need to be fully recognized or associated with serious burden. Instead, the starting point for prevention should be that mild mental health complaints might easily deteriorate in the (near) future, with consequent risks for the well-being, health, and safety of the worker and his or her patients. This message should be the basis for the discussion of screening results with the worker and for the initiation of advises the occupational physician provides. The training should also include the discussion of any expected barriers to the initiation of more intensive advising as part of prevention.

In addition to the need for more proactive guidance of the group of workers with mild or moderate severe complaints, more proactive guidance by the occupational physicians for a small selected group of workers with more severe mental health complaints is regarded as necessary. In previous studies where screening and intervention were combined to improve work outcomes and the mental health states of depressive workers, considerable effects on depressive symptoms[8 14] and work functioning were observed.[15] In all three studies, screening targeted more severe mental health complaints and the treatment offered was much more intensive, in contrast to the WHS mental module studied. Therefore, for this group of workers, the occupational physician should organize more intensive counseling by referral to specialized mental health caregivers, such as psychologists and welfare workers. Referral to an employee assistance program might also be useful,[16 17] if these are included in the national or company occupational health care program, as is the case in the United Kingdom, Japan, and the United States. The process evaluation showed that ten workers were referred to other caregivers; however, no increase in use of those care-givers is found at three months of follow-up. Thus, it can be questioned whether the care was actually set-up by the workers. For that reason, after referral, the occupational physician should follow-up on these workers to verify that the needed help has been scheduled and to provide possible guidance for more work-related regulations.

With more proactive effort during the preventive consultation, the worker's mental health and work functioning might be enhanced. Nevertheless, regarding the absence of a positive effect on mental health complaints, one must keep in mind that the causes of mental health complaints may be situated in the work environment. The improvement of (mild) mental health complaints by interventions targeting the individual might be difficult as long as high job-demands and unfavorable work environment characteristics remain unchanged.

In conclusion, the WHS mental module is a successful strategy in stimulating seeking help from the occupational physician. Regarding the aims of the WHS formulated by the Netherlands Society of Occupational Medicine,[18] we can conclude that the WHS mental module studied here does monitor and improve work functioning. However, the goals of monitoring and improving mental health were not sufficiently met by the studied WHS mental module during the six-months follow-up period.

RECOMMENDATIONS

Improvement of the workers' health surveillance mental module

Screening

- It is recommended to tune the selection of cut-off values to the feedback strategy for the individuals' screening results. When providing feedback personally by the occupational

physicians, a decision to aim for high sensitivity but low specificity is tolerable. When choosing to present personal screening results online, cut-off values with higher specificity and lower sensitivity should be chosen.

Preventive consultation with the occupational physicians

- It is suggested to improve the method of invitation to the preventive consultation. As workers prefer to make appointments themselves, this change in strategy might reduce reluctance in workers to attend the preventive consultation.
- It is advised to better train the occupational physicians for the preventive context. Aspects that should be more intensively discussed in their training include communication with workers with mild mental health complaints rather than an expressed need for help.

Future research

Further evaluation of the NWFQ

- The clinimetric quality of subscale 2), *Impaired decision making*, was too poor to recommend further use of this subscale in the present form. Impaired decision making is still regarded as an important aspect of the construct of the NWFQ. Future studies should improve the current items and response categories or identify suitable new items, to yield a reproducible and valid form of this subscale that can be included in the NWFQ.
- Cut-off values of the NWFQ should be further developed. A method analogous to the anchor-based method for the MIC-value identification or a Delphi consensus study might be suitable to identify cut-off values for the NWFQ.
- The NWFQ was developed and evaluated in the Dutch language. Backward-forward translations of the questionnaire exist in English and in Italian. Future research should evaluate the clinimetric quality of these versions.
- Future studies should evaluate the validity of the NWFQ for related occupational groups such as caregivers outside the hospital or physicians.

Future research on workers' health surveillance mental modules

- The effect of providing feedback on screening results and of the preventive consultation with the occupational physician on the recognition of mental health complaints should be further investigated.
- More insight is needed into the advice provided by occupational physicians to address impaired work functioning as well as into the actions workers initiate to improve their work functioning. With more transparency in terms of the advice given and followed, the preventive consultation protocol could be improved, which might lead to an increased effect on work functioning.

Implications for practice

Policy makers

- The work functioning impairments in nurses and allied health professionals due to CMDs are associated with risks for the patient's well-being, health, and safety. Policy makers in medical centers should be aware of these risks. Making use of a WHS mental module or similar strategies, such as employee assistance programs, can help reduce work functioning impairments due to CMDs in this group.

Occupational health professionals

- Actions to further develop and evaluate WHS mental modules, within the healthcare service as well as in other sectors, should be initiated.
- If workers are noted to have mental health problems, although mild, occupational health professionals should always discuss possible consequences with regard to work functioning, as well as the risk for deterioration of mental health complaints in the (near) future.

REFERENCES

1 Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychol Assessment* 1995;7:238-47.
2 Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
3 Lerner D, Amick Iii BC, Rogers WH et al. The Work Limitations Questionnaire. *Med Care* 2001;39:72-85.
4 Schaufeli WB, Vandierendonck D. A Cautionary Note About the Cross-National and Clinical Validity of Cutoff Points for the Maslach Burnout Inventory. *Psychol Rep* 1995;76:1083-90.
5 Schaufeli WB, Bakker AB, Hoogduin K et al. On the clinical validity of the Maslach Burnout Inventory and the Burnout Measure. *Psychology & Health* 2001;16:565-82.
6 Rutjes AW, Reitsma JB, Coomarasamy A et al. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007;11:iii, ix-51.
7 Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995;311:376-80.
8 Lexis MA, Jansen NW, Huibers MJ et al. Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial. *Occup Environ Med* 2011;68:400-7.
9 Schene AH, Koeter MW, Kikkert MJ et al. Adjuvant occupational therapy for work-related major depression works: randomized trial including economic evaluation. *Psychol Med* 2007;37:351-62.
10 van der Klink JJ, Blonk RW, Schene AH et al. Reducing long term sickness absence by an activating intervention in adjustment disorders: a cluster randomised controlled design. *Occup Environ Med* 2003;60:429-37.
11 Brouwer WB, van Exel NJ, Koopmanschap MA et al. Productivity costs before and after absence from work: as important as common? *Health Policy* 2002;61:173-87.
12 Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health* 2000;54:502-9.
13 Nygaard P, Aasland OG. Barriers to implementing screening and brief interventions in general practice: findings from a qualitative study in Norway. *Alcohol* 2011;46:52-60.
14 Wang PS, Simon GE, Avorn J et al. Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. *JAMA* 2007;298:1401-11.
15 Lerner D, Adler D, Hermann RC et al. Impact of a Work-Focused Intervention on the Productivity and Symptoms of Employees With Depression. *J Occup Environ Med* 2012;54:128-35.
16 Merrick EL, Hodgkin D, Hiatt D et al. Patterns of service use in two types of managed behavioral health care plans. *Psychiatr Serv* 2010;61:86-9.
17 Mains DA, Fairchild TJ, Rene AA. An employee assistance program for caregiver support. *J Gerontol Soc Work* 2006;47:157-73.
18 Netherlands Society of Occupational Medicine (NVAB). Guidance Workers' Health Surveillance [In Dutch: Leidraad Preventief Medisch Onderzoek]. Utrecht, the Netherlands: NVAB, 2005.

Summary

Common mental disorders (CMDs) can lead to impairments in work functioning. In addition to financial consequences, in certain occupations, impairments in work functioning can have serious consequences for other individuals. In the work of nurses and allied health professionals, impaired work functioning due to CMDs is related to higher risks for the workers' and patients' health and safety, e.g., by increasing the frequency of medication errors and needle stick injuries, and by decreasing patient well-being. Both the high prevalence of CMDs and the risk of serious adverse events caused by CMDs in the work of nurses and allied health professionals necessitate a call for research and presumably action on this topic.

If we are able to identify concrete aspects of work functioning impairments due to CMDs, future research on the prospects for both early detection of employees with CMDs and work functioning impairments and purposeful interventions will be possible. Therefore, the first aim of this thesis was to develop and evaluate a job-specific instrument to assess work functioning impairments due to common mental disorders in nurses and allied health professionals. A conceptual model was developed for the specification of work functioning aspects. In this model, work functioning is composed of two main elements, task and contextual work functioning, and it encompasses four dimensions: process, quality outcomes, quantity outcomes, and effort.

Furthermore, preventive actions are necessary to promote good mental health and work functioning in nurses and allied health professionals. Workers' health surveillance (WHS) is a possible strategy for preventive action by occupational health services. Therefore, the second aim of this thesis was to develop and evaluate a mental module for WHS for nurses and allied health professionals. The WHS mental module aims to use an online screening to identify workers with mild to moderate severe symptoms of CMDs and/or early signs for impaired work functioning. It is hypothesized that workers who receive the WHS mental module will seek more help compared to the control group. Furthermore, it is expected that work functioning impairments and mental health complaints will improve more for this group of workers.

The literature review presented in **Chapter 2** aimed to inventory aspects of the work functioning of nurses and allied health professionals that are affected by CMDs. A sensitive, systematic literature search was conducted using four electronic databases: PubMed, PsycINFO, Embase, and Cinahl. All studies published between 1998 and 2008 that examined a relationship between CMDs and work functioning in nurses or allied health professionals were included. The data from these investigations were categorized into themes, and the strength of the evidence for each theme was assessed. The search resulted in 2,792 studies, of which 16 met the inclusion criteria. Of these, 13 had a cross-sectional design, one was a vignette study, and two were narrative reviews. In all of the studies, the subjects were nurses. The retrieved aspects of sub-optimal work functioning due to common mental disorders were merged into 15 themes. Strong evidence was found for five themes: general errors,

medication errors, near misses, patient safety, and patient satisfaction. Moderate evidence was found that CMDs are associated with complex motor skills and with general performance, whereas evidence for an association between common mental disorders and needle stick injuries was inconclusive. Seven themes were only supported by narrative evidence: interpersonal behavior, energy, focus on goals and responsibility, work speed, avoiding work while on the job, coping with emotions, and motivation.

The development of the job-specific questionnaire for work functioning impairments due to CMDs was described in **Chapter 3**. As a first step, an item pool was developed. In addition to the results of the literature review, five focus groups were held with employees and experts to identify additional themes and concrete signs in the characteristic behaviors of impaired work functioning in nurses and allied health professionals with CMDs. Based on the results from the literature study and focus groups, 13 themes were formulated representing different aspects of work functioning impairments which are supposed to be assessable by self-report. For each theme, items were phrased representing examples of work behavior and tasks that might be impaired. In six verbal probe interviews, the items and the underlying themes were checked for its relevance and completeness and rephrased if necessary. This process resulted in an item pool of 14 themes and a total of 231 items.

In a second step, definite subscales and items were identified. In a cross-sectional assessment, 314 nurses and allied health professionals filled out the item pool. In explorative factor analysis (using Principal Component Analysis), the subscales' structure was generated, and definite items were chosen. For a corroboration of the subscales, confirmative factor analysis was performed (using the Oblique Multiple Group Method). The study resulted in a questionnaire with seven subscales and a total of 50 items, called the Nurses Work Functioning Questionnaire (NWFQ). The seven subscales are the following: 1) *Cognitive aspects of task execution and general incidents*, 2) *Impaired decision making*, 3) *Causing incidents at work* (not suitable for allied health professionals), 4) *Avoidance behavior*, 5) *Conflicts and irritations with colleagues*, 6) *Impaired contact with patients and their family*, and 7) *Lack of energy and motivation*. The internal consistency was good for four subscales (Cronbach's alphas ≥ 0.80) and acceptable for three subscales (Cronbach's alphas ≥ 0.70). To evaluate the content validity, six experts evaluated the subscales and items on their relevance and representativeness. All of the aspects were evaluated as relevant or highly relevant, and the representativeness of the item pool was assessed as highly comprehensive. The finding of a variety of subscales supports the idea that a multidimensional approach is suitable for the study of impaired work functioning and emphasizes the added value of a job-specific approach.

The clinimetric quality of the newly developed NWFQ was further studied in **Chapter 4** by evaluating the reproducibility and construct validity of the NWFQ. To achieve this aim, the

questionnaire was administered to 314 nurses and allied health professionals with a re-test in 112 of the subjects after two weeks.

To evaluate the reproducibility of the NWFQ, the reliability and the level of agreement of the questionnaire were assessed. The reliability – the ability of a test to distinguish between persons despite the measurement error – was evaluated by the intraclass correlation coefficient (ICC). The level of agreement represents the extent of the expected difference between repeated measures, due to measurement error, which is expressed by the standard error of measurement (SEM). All of the NWFQ subscales showed good reliability, with ICC values above 0.70, except for subscale 2), *Impaired decision making*, for which the ICC was too low (ICC = 0.16). The standard error of measurement ranged from 3 to 6 for six subscales, which is considered to be low for a scale ranging from 0 to 100. However, for subscale 2), the standard error of measurement was rather high (SEM = 17).

Three aspects of construct validity were studied: convergent validity, divergent validity, and discriminative validity. Medium (> 0.30) to high (> 0.60) correlations were hypothesized between the NWFQ scores and those of a general work functioning scale, the Endicott Work Productivity Scale (convergent validity), and low (< 0.30) correlations were expected between the NWFQ scores and those of a physical functioning subscale of the SF-36 (divergent validity). For discriminative validity, a statistically significant difference between the NWFQ results of groups of subjects with and without mental health complaints was hypothesized, which was analyzed by a Mann Whitney U test.

Our data offer strong support for good construct validity, with the exception of subscale 2), *Impaired decision making*. Regarding the convergent validity of these six subscales, all correlations were substantial and in line with the hypothesis. The fact that these correlations were medium and not high affirms that the NWFQ has significant overlap with a generic work functioning scale, yet is an instrument that possesses additional value because it measures job-specific aspects that may be neglected by a generic questionnaire. The hypothesis for good divergent validity was supported for all of the seven subscales. Additionally, all of the subscales showed good discriminative validity. Therefore, the relatedness of CMDs to impaired work functioning is evident.

In conclusion, the NWFQ demonstrated good clinimetric properties for six subscales. Subscale 2), *Impaired decision making*, did not show enough ability to discriminate between subjects, and its association with other work functioning measures was too weak. Therefore, we discourage the use of this subscale in its present form. In the Appendix of this thesis the NWFQ is presented.

Chapter 5 focused on the evaluation of another measurement characteristic for the NWFQ, the interpretability of change. The aims of this study were, first, to identify the minimal important change values for improvement and the smallest detectable change values and second, to evaluate the interpretability of individual change scores. The analyses were

conducted both for the overall NWFQ score and separately for six subscales. The subscale *Impaired decision making* was excluded from this analysis. Knowledge of the smallest detectable change and the minimal important change helps interpret the change scores of the Nurses' Work Functioning Questionnaire on an individual level. These values might guide researchers and practitioners in their conclusions regarding whether changes in individual workers are *real* and *relevant*.

Data collected in a randomized controlled trial at two distinct time points, baseline and three months of follow-up, were used to achieve the aforementioned aims. In this trial, 358 nurses and allied health professionals at one Dutch academic medical center participated at both measurement points. The minimal important change values were calculated using two anchor-based methods, the mean-change method and the ROC-curve method. Additionally to the MIC calculation based on absolute change scores, MIC values based on relative change scores were calculated, to correct for the baseline scores. The smallest detectable change values were calculated using the standard error of measurement. Requirements for good interpretability were two-fold: first, MIC values had to be higher than the SDC values and second, the area under the curve (AUC) values of the ROC-curve had to be 0.70 or higher. The interpretation was based on results from a subgroup of workers who demonstrated work functioning impairments at the first measurement point (high baseline scores on the NWFQ), as only in this subgroup was improvement realistic and assessable by the NWFQ.

The minimal important change values ranged from 4.4 to 29 for the mean-change method and from 9.5 to 41.5 for the ROC-curve method. The smallest detectable change values ranged from 7 to 17. These were considered to be small compared with the scale range (0 to 100), except for two values. Regarding the interpretability, 10 of the 14 calculated minimal important change values exceeded the smallest detectable change values in the subsample with high baseline scores and three AUC values were of ≥ 0.70. Thus, the requirements for a good interpretability of change were met in three of the seven tested scales: *Causing incidents at work, Impaired contact with patients and their family,* and *Lack of energy and motivation.*

The calculations in this study were based on a small number of subjects because only a few workers in this study reported that they perceived an important improvement in their work functioning between the two time points examined. The replication of the minimal important change value calculations in a population in which the occurrence of important improvements in work functioning is expected to be higher is therefore recommended. Another concern to address is that of an applicable anchor question. In this study, the formulation of an accurate anchor question for perceived change in work functioning appeared to be complex. Therefore, further research is regarded as necessary to identify a more suitable formulation for anchor questions on perceived change in work functioning. Provisionally, the calculated minimal important change values can cautiously be used.

The second aim of this thesis was to develop and evaluate a workers' health surveillance (WHS) mental module, which is expected to be an effective preventive strategy to monitor and promote good (mental) health and work functioning. **Chapter 6** describes the protocol for a study in which the effectiveness of two strategies for a WHS mental module were evaluated. One strategy included the care of an occupational physician (WHS OP-care). The second strategy included the choice of self-help e-mental health interventions for the workers (WHS EMH-care). The study evaluating the WHS mental module was designed as a cluster randomized controlled trial consisting of three arms (two intervention groups and one control group), with randomization at the ward level. The study population consisted of 86 departments in one Dutch academic medical center and includes a total of 1731 nurses and allied health professionals. The effectiveness of the WHS OP-care strategy compared with a control group was a subject of this thesis.

The first step of the WHS OP-care mental module included an online screening for work functioning impairments and six types of mental health complaints: distress, the need for recovery, risky drinking behavior, depression, anxiety, and post-traumatic stress disorder. Workers who participate in the screening immediately received feedback regarding their personal results online. In the second step, workers in the WHS OP-care group who are screened as positive received an invitation for a consultation with their occupational physician. In the control group, workers received no screening results or intervention. It was expected that, in the group of workers who received the WHS mental module, the WHS OP-care group, more workers would seek help than the control group. A reduction in work functioning impairments and mental health complaints was also expected in the WHS OP-care group compared with a control group. These effects were measured three and six months after the baseline.

In **Chapter 7**, the effects of the WHS OP-care strategy for a workers' health surveillance mental module, which includes screening plus feedback and an invitation for a consultation with the occupational physician, if applicable, are presented and compared with the results of a control group. In total, 1,152 workers from 57 wards were invited to participate in the WHS OP-care group or the control group. Of these workers, 191 nurses and allied health professionals took part in the WHS OP-care group at baseline. In the control group, 188 workers completed the baseline questionnaire. A total of 151 workers (79%) in the WHS OP-care group were classified as having work functioning impairments and/or mental health complaints, compared with 161 (86%) in the control group. A statistically significant interaction effect was found for study-group*time on help-seeking behavior ($p = 0.02$), which indicates that the time course of help-seeking behavior differed in a statistically significant way between the two study groups. After the baseline, help-seeking behavior decreased in the control group, and stabilized between three and six months after the initial screening. In the WHS OP-care group, the percentage of participants engaging in help-seeking behavior was

approximately stable between baseline and three months of follow-up and was therefore higher than that of the control group. The difference in this percentage between the groups was 13%, which was marginally significant (p = 0.09). At the six-month follow-up, the help-seeking behavior frequency decreased more steeply in the WHS OP-care group than it did in the control group until reaching a level lower than that observed in the control group (8% lower), however, without a statistically significant discrepancy in the percentages between the control and WHS OP-care group (p = 0.14).

Workers who received the WHS mental module demonstrated less work functioning impairments in the follow-up measurements at three and six months after the baseline (p = 0.04). At three months of follow-up, 45% of the workers who received the intervention showed at least minimal important improvements in work functioning; this was 30% in the control group (p = 0.03). At the six-month follow-up, 41% of workers in the WHS OP-care group exhibited improvements representing a change that was minimal or larger; the magnitude of this change in the control group was 28%. This difference was marginal significant (p = 0.05). No statistically significant effects were found on mental health complaints, except for an interaction effect of study-group*time for risky drinking behavior (p < 0.01). The difference between the two study groups regarding risky drinking behavior was statistically significant at neither the three-month follow-up (p = 0.08) nor at the six-month follow-up (p = 0.36), but the time course for drinking behavior differed. At three months of follow-up, the WHS OP-care group showed a slight rise in drinking behavior compared with the control group; however, at six months of follow-up, risky drinking behavior in the WHS OP-care group decreased to below the levels observed in the control group.

In conclusion, the WHS mental module is regarded as an effective strategy to stimulate seeking help from the occupational physician and in improving work functioning in workers with mild to moderate severe work functioning and/or mental health complaints. It would be interesting to further investigate the possible effects of the WHS mental module on the recognition of mental health complaints.

Chapter 8 presents the process evaluation of the randomized controlled trial in which the effectiveness of the WHS mental module is studied. This chapter aimed to study the participants' response to and compliance with the intervention and the occupational physicians' adherence to the consultation protocol, as well as to describe the perspectives of the participants and occupational physicians on the WHS mental module.

The process evaluation of the WHS mental module revealed that the response rate to the intervention offered in this study was 32%. Fifty-one (34%) of the 151 workers who were screened as positive, visited the occupational physician for a preventive consultation. Due to system errors, only 125 of the 151 workers were invited for a consultation, thus 41% of the invited workers followed up upon the invitation. Of these workers 80% followed the occupational physician's advice if such advice was given. The occupational physicians

reported that 70% of the participants who went to the consultation felt that the personal results (partially) reflected their perceived work functioning, whereas 60% felt similarly for the mental health screening results. Moreover, almost all (97%) of the workers who visited the occupational physician for the preventive consultation felt that they could be open and honest with their occupational physicians. The preventive consultation was perceived as being effective by nine out of 15 participants. Most of the participants would appreciate being offered a WHS mental module in the future.

The adherence of the occupational physicians to the consultation protocol was high, as in most cases, all of the seven steps were followed. The occupational physicians were satisfied with the consultation protocol and training. They felt that the preventive consultations had been meaningful and that implementing a WHS mental module in the future would also be meaningful. Suggestions were given for improvements of the studied strategy for a WHS mental module, e.g., regarding the in-house communication strategy and the role of e-mental health interventions. The occupational physicians expressed some reluctance and helplessness regarding providing advice and initiating further care in cases when the worker did not recognize the results of the mental health screening.

In conclusion, the WHS OP-care strategy was well-received by the nurses and allied health professionals who participated in the study. Strategies to improve the response of workers and the protocol adherence of the occupational physicians should be further investigated.

In **Chapter 9**, the main research findings were summarized and discussed, and recommendations for further research and practice are presented. It is concluded that the aim of developing a job-specific work functioning questionnaire for nurses and allied health professionals with CMDs has been successfully achieved by the creation of the Nurses' Work Functioning Questionnaire (NWFQ). With its seven subscales, six of which demonstrated good clinimetric properties, the NWFQ demonstrates both breadth and depth of measurement and permits self-administration. The individual subscale scores provide insight into the precise aspects of impaired work functioning, allowing for the tailoring of interventions to individual needs.

For future research, further improvements of the subscale *Impaired decision making* are recommended. In addition, future research should focus on the further validation of cut-off values to identify workers with work functioning impairments.

In conclusion, the WHS mental module is a successful strategy in stimulating seeking help from the occupational physician. Regarding the aims of WHS formulated by the Netherlands Society of Occupational Medicine, it was concluded that the WHS mental module studied here does monitor and improve work functioning. However, the goals of monitoring and improving mental health were not sufficiently met by the studied WHS mental module. Improvements in the screening strategy are recommended, such as by attuning the cut-off values of the

feedback strategy. Additionally, in the future, occupational physicians should be better trained for preventive care.

Practical implications were given in this study, including the recommendation that policy makers in hospitals should be aware of the risks posed by CMDs and related work functioning impairments in healthcare workers. Additionally, employers should take responsibility for the health of their workers and for the well-being and safety of the hospitals' patients by making use of workers' health surveillance mental module. Occupational health services should undertake actions to further develop, evaluate, and stimulate mental health modules as part of workers' health surveillance, not only within healthcare service but also in other sectors.

Samenvatting

Aanleiding

Psychische klachten zijn een veel voorkomend probleem onder de beroepsbevolking in Nederland. Het gaat daarbij om spannings- en stressklachten, psychische vermoeidheid, depressieve en angstklachten, maar ook overmatig alcohol gebruik. In een periode van 12 maanden heeft in Nederland circa 9% van de beroepsbevolking te maken met depressie, angstklachten of alcoholproblemen. Voor stressklachten, milde depressieve of angstklachten is dit percentage hoger, tussen de 19% en 26%. De gevolgen van een verminderde psychische gezondheid blijven niet beperkt tot individueel lijden. Psychische klachten – ook lichte klachten – kunnen een belemmering vormen voor optimaal functioneren tijdens het werk. Dit uit zich in verzuim, maar ook in minder goed functioneren van werkenden die ondanks gezondheidsproblemen toch aan het werk zijn. Bekende gevolgen van psychische klachten zijn bijvoorbeeld een verminderde concentratie, het maken van fouten, de extra inspanning die nodig is om het werk uit te voeren en verminderd interpersoonlijk contact. Het is echter nog niet duidelijk welke aspecten van het werk precies door psychische klachten worden beïnvloed.

In de zorgsector hebben werknemers een hoog risico op het ontstaan van psychische klachten. Dit is met name het geval onder verpleegkundigen. Hiervoor zijn diverse oorzaken aan te wijzen, waaronder de hoge werkdruk, zware emotionele belasting en blootstelling aan ongewenst gedrag door patiënten of bezoekers. Tegelijkertijd kan verminderd functioneren tijdens het werk juist bij deze groep ernstige gevolgen hebben. Zo kan het risico op fouten toenemen, wat de patiëntveiligheid in gevaar brengt. Er kan dan ook gesproken worden van een dubbel risico voor werknemers in de gezondheidszorg: het hoge risico op het ontstaan van psychische klachten met als gevolg suboptimaal functioneren, én het risico op negatieve gevolgen voor de patiënten. Gezien het aantal mensen dat werkzaam is in de gezondheidszorg, is dit een omvangrijk probleem. Inzicht verkrijgen in het effect van psychische gezondheidsklachten op het functioneren van verpleegkundigen en paramedici is daarom een van de centrale doelen in dit proefschrift.

Voor gericht onderzoek naar dit vraagstuk is meer inzicht in de precieze aspecten van verminderd werkfunctioneren ten gevolge van psychische klachten noodzakelijk. Er is behoefte aan onderzoek naar interventies die verbetering kunnen brengen in verminderd werkfunctioneren door psychische klachten. Voor dergelijk onderzoek zijn meetinstrumenten nodig die verminderd werkfunctioneren in kaart kunnen brengen. Hoewel er valide vragenlijsten bestaan die diverse psychische klachten meten, is er tot op heden nog geen vragenlijst ontwikkeld die de aan psychische klachten gekoppelde beperkingen in het functioneren van verpleegkundigen en paramedici in kaart kan brengen. Dit proefschrift richt zich dan ook op de ontwikkeling, validering en evaluatie van een dergelijke vragenlijst. Gekozen is voor de vorm van een zelfrapportage vragenlijst.

Indien psychische klachten onbehandeld blijven, kunnen deze lang aanhouden en verergeren. Vaak zoeken werknemers met psychische klachten echter pas laat hulp, bijvoorbeeld pas na een ziekmelding of na ongelukken of incidenten in het werk. Een actief preventiebeleid voor werknemers is daarom belangrijk. Door werknemers met psychische klachten en daaraan verbonden functioneringsproblemen al in een vroeg stadium op te sporen, zou bij deze groep de zelfreflectie over klachten en functioneren bevorderd kunnen worden. Hiermee kan het zoeken naar adequate hulp worden gestimuleerd. Een methode voor een dergelijke preventiestrategie binnen het (arbo)zorgsysteem is een preventief medisch onderzoek (PMO). PMO heeft als uitgangspunt effectieve preventie, met als doel het bewaken en bevorderen van de gezondheid en het functioneren van werknemers. In een PMO kunnen na screening doelgericht interventies aan werknemers worden aangeboden. Er is echter nog weinig onderzoek gedaan naar de effecten van PMOs die zich specifiek richten op psychische gezondheidsaspecten. Dit proefschrift richt zich daarom op de evaluatie en ontwikkeling van een psychische module PMO voor verpleegkundigen en paramedici.

Samenvattend heeft dit proefschrift twee doelen:

Doel 1: Het ontwikkelen en valideren van een vragenlijst voor het meten van verminderd functioneren in het werk bij psychische klachten van verpleegkundigen en paramedici.

Doel 2: Het ontwikkelen en evalueren van een psychische module PMO voor verpleegkundigen en paramedici. De verwachting is dat werknemers met psychische gezondheidsklachten en/of verminderd werkfunctioneren eerder hulp zullen zoeken door deelname aan een psychische module PMO. Tevens wordt verwacht dat door gerichte suggesties voor oplossingen door de hulpverlener, het werkfunctioneren en psychische gezondheidsklachten zullen verbeteren.

Ten aanzien van deze twee doelen werden vijf onderzoeksvragen geformuleerd. Hieronder worden de studies van dit proefschrift samengevat aan de hand van deze vijf onderzoeksvragen.

Welke aspecten van het werk van verpleegkundigen en paramedici worden aangetast door psychische klachten? (Hoofdstukken 2 en 3)

Voor het ontwikkelen van een vragenlijst is eerst onderzocht welke aspecten van het werk van verpleegkundigen en paramedici negatief beïnvloed worden door psychische klachten. Om een zo compleet mogelijke lijst met aspecten in kaart te brengen, is een systematisch literatuuronderzoek uitgevoerd en zijn vijf focusgroepen gehouden met werknemers en met experts op het gebied van arbeid en gezondheid.

Samen leverden de zestien gevonden studies uit het literatuuronderzoek en de analyse van de focusgroepen dertien aspecten op van verminderd functioneren door psychische klachten, die met zelfrapportage te meten zijn. Hierbij gaat het om 1) taakgerichte aspecten, zoals het overzicht verliezen, verminderd geheugen of verminderde besluitvorming; 2) het veroorzaken van fouten of bijna-fouten in het werk; 3) interpersoonlijke aspecten, zoals verminderd contact met patiënten, vermijden van contact met collega's of agressief gedrag vertonen en 4) intrapersoonlijke aspecten van het werk, zoals het werk als zwaarder ervaren of verminderde motivatie.

Hoe ziet een vragenlijst eruit die verminderd functioneren in het werk van verpleegkundigen en paramedici met psychische klachten meet? Wat is de inhoudsvaliditeit en welke subschalen kunnen worden onderscheiden? (Hoofdstuk 3)

De eerste stap bij de ontwikkeling van de vragenlijst was het samenstellen van een lijst met mogelijke vragen, op basis van de dertien aspecten die naar voren kwamen in de literatuurstudie en in de focusgroepen. Voor elk van deze aspecten zijn mogelijke vragen geformuleerd. De dertien aspecten en bijbehorende vragen zijn vervolgens aan zes experts voorgelegd, ter beoordeling van de relevantie en de volledigheid. Hierna zijn de vragen, waar nodig, aangepast. Door zes verpleegkundigen en paramedici zijn de vragen geëvalueerd op duidelijkheid. Dit resulteerde in een lijst met 213 mogelijke vragen.

In de tweede stap is uit de lijst met mogelijke vragen na statistische analyse de uiteindelijke vragenlijst samengesteld. Via een online enquête is de lijst met mogelijke vragen door 314 verpleegkundigen en paramedici ingevuld. Door middel van factoranalyse werden overkoepelende thema's voor de vragen geïdentificeerd (exploratieve factoranalyse) en bevestigd (confirmatieve factoranalyse). Aan de hand hiervan werd uiteindelijk de definitieve vragenlijst vastgesteld: de Nurses Work Functioning Questionnaire. Deze vragenlijst omvat 50 vragen, verdeeld over zeven thema's (subschalen genoemd). De zeven subschalen zijn: 1) *Cognitieve aspecten van taakuitvoering en fouten in het algemeen*, 2) *Moeite met besluitvorming*, 3) *Veroorzaken van fouten in het werk* (deze subschaal is niet geschikt voor paramedici), 4) *Vermijdingsgedrag* 5) *Conflicten en irritaties met collega's* 6) *Verminderd contact met patiënten en hun familie*, 7) *Gebrek aan energie en motivatie*. In de Appendix van dit proefschrift is de uiteindelijke vragenlijst opgenomen.

Op basis van zes expertchecks kon geconcludeerd worden dat de vragenlijst een goede inhoudsvaliditeit heeft. Alle aspecten van de ontwikkelde vragenlijst werden als relevant of uitermate relevant beoordeeld en in zijn geheel werden de vragenlijst en de subschalen compleet bevonden.

De inhoudelijke samenhang van de vragen die bij eenzelfde subschaal horen, wordt gemeten aan de hand van de interne consistentie. Deze wordt uitgedrukt door middel van de Cronbach's alpha, waarbij een waarde boven 0.80 goed is tussen 0.70 en 0.80 acceptabel. De interne consistentie was voor vier subschalen goed en voor drie subschalen acceptabel.

Wat zijn de klinimetrische eigenschappen van de ontwikkelde vragenlijst, de Nurses Work Functioning Questionnaire (NWFQ), in termen van reproduceerbaarheid, constructvaliditeit en interpreteerbaarheid van verandering? (Hoofdstukken 4 en 5)

Voor een verantwoord gebruik van de vragenlijst is de NWFQ verder onderzocht op drie klinimetrische eigenschappen: de reproduceerbaarheid (Hoofdstuk 4), de constructvaliditeit (Hoofdstuk 4) en de interpreteerbaarheid van veranderingen (Hoofdstuk 5). Deze eigenschappen zijn apart onderzocht voor de zeven subschalen van de NWFQ.

De reproduceerbaarheid geeft inzicht in de betrouwbaarheid en de overeenkomst tussen scores bij herhaalde metingen bij dezelfde personen. Ter beoordeling hiervan hebben 112 verpleegkundigen en paramedici de vragenlijst tweemaal ingevuld binnen een periode van twee weken. De betrouwbaarheid werd vastgesteld door middel van de intraclass correlatie coëfficiënt (ICC); deze loopt van 0 tot 1, waarbij een waarde boven 0.70 wenselijk is. Voor zes subschalen was de betrouwbaarheid goed (ICC tussen 0.72 en 0.88), voor subschaal 2), *Moeite met besluitvorming*, was de betrouwbaarheid te laag (ICC = 0.16). Om de overeenkomst tussen scores bij herhaalde metingen uit te drukken is de standaard meetfout berekend voor de NWFQ subschalen. Deze dient zo laag mogelijk te zijn. De standaard meetfout was laag tot matig voor zes subschalen (SEM: 2.95 - 6.12) en hoog voor subschaal 2), *Moeite met besluitvorming*, (SEM = 17.11). Samenvattend is de reproduceerbaarheid voor zes van de zeven subschalen goed. Voor voor subschaal 2), *Moeite met besluitvorming*, is de reproduceerbaarheid onvoldoende.

De constructvaliditeit is bedoeld om inzicht te geven of de ontwikkelde vragenlijst meet wat deze beoogt te meten. Hiertoe hebben 314 verpleegkundigen en paramedici de NWFQ ingevuld, evenals vier aanvullende vragenlijsten waarmee de NWFQ scores vergeleken werden ter beoordeling van de constructvaliditeit. Zes van de zeven subschalen vertonen een goede overeenkomst met een vragenlijst die algemeen werkfunctioneren meet; de correlatie coëfficiënten waren matig tot hoog (tussen 0.38 en 0.63). Voor subschaal 2), *Moeite met besluitvorming*, was de samenhang met een vergelijkbare vragenlijst echter onvoldoende; de correlatie coëfficiënt was laag (0.22). Alle zeven subschalen vertonen, zoals verwacht, een lage overeenkomst met een vragenlijst voor fysiek functioneren; de correlatie coëfficiënten waren laag (tussen -0.06 en -0.23). Bovendien laten alle zeven subschalen een verwacht verschil zien tussen een groep werknemers met psychische klachten en een groep werknemers zonder psychische klachten. Het is belangrijk om te benadrukken dat hiermee de onderliggende samenhang tussen verminderd werkfunctioneren en psychische klachten is aangetoond voor de NWFQ. Samenvattend is voor zes subschalen de constructvaliditeit als goed beoordeeld. Dit geldt alleen niet voor subschaal 2), *Moeite met besluitvorming*.

Naar aanleiding van dit onderzoek is besloten om subschaal 2), *Moeite met besluitvorming*, in zijn huidige vorm niet te gebruiken. Wel wordt moeite met besluitvorming

als een belangrijk onderdeel beschouwd van verminderd werkfunctioneren ten gevolge van psychische klachten. Het werd in de ontwikkelingsfase van de vragenlijst herhaaldelijk genoemd in de focusgroepen en werd door experts als relevant beoordeeld. Toekomstig onderzoek zou zich moeten richten op een verbetering van deze subschaal, bijvoorbeeld door het herformuleren van vragen, het toevoegen van vragen of het aanpassen van de antwoordcategorieën.

Het doel van de studie naar de interpreteerbaarheid van veranderingen in individuen op scores van de NWFQ was tweeledig. Het eerste doel was de bepaling van normscores voor een minimaal belangrijke verbetering (minimal important change - MIC). Dit zijn verbeteringen in scores van werkfunctioneren waarvan gesteld kan worden dat deze daadwerkelijk van betekenis zijn voor de werknemer. Ook werden normscores vastgesteld voor de kleinst mogelijke verandering (smallest detectable change - SDC) die de NWFQ kan meten. Dit is een verschilscore waarbij met hoge zekerheid (95%) gesteld kan worden dat deze niet op de standaard meetfout berust. Het tweede doel was om te beoordelen of interpretatie van verschilscores met de NWFQ mogelijk is. Deze analyses zijn uitgevoerd voor de totale score van de NWFQ en voor zes afzonderlijke subschalen; subschaal 2), *Moeite met besluitvorming*, is hierin niet meegenomen.

Voor deze doelen is bij 358 verpleegkundigen en paramedici op twee momenten de NWFQ afgenomen, op een eerste meetmoment en drie maanden later. Daarnaast is op het tweede meetmoment ook een eigen inschatting gevraagd over de verandering van het eigen werkfunctioneren. Deze eigen inschatting werd gebruikt om de MIC te bepalen. De SDC is bepaald door middel van de standaard meetfout (SEM) die uit het vorige onderzoek bekend was. Omdat deze studie gaat over scores omtrent verbetering in werkfunctioneren, zijn de resultaten gebaseerd op analyses in een subgroep van werknemers die op het eerste meetmoment problemen in hun werkfunctioneren vertoonden. Dit omdat alleen in deze groep een aanzienlijke verbetering verwacht kon worden.

Voor de bepaling van de MIC waarden zijn twee methoden gehanteerd. Op basis van de ‘mean-change method’ varieerden de MIC waarden van 4 tot 29 Op basis van de ‘ROC-curve method’ varieerden de MIC waarden van 10 tot 42. De SDC was voor vier subschalen goed; de waarden lagen tussen 7 en 12, wat voor een vragenlijst met scores van 0 - 100 als laag beschouwd kan worden. Voor drie subschalen was de SDC hoog, met waarden van rond de 17. Geconcludeerd wordt dat drie subschalen van de NWFQ voldoen aan de eisen voor goede interpreteerbaarheid van verandering: *veroorzaken van incidenten in het werk, verminderd contact met patiënten en hun familie, en gebrek aan energie en motivatie.* Voor de overige subschalen kan een goede interpreteerbaarheid nog niet worden aangetoond.

Omdat de steekproef voor dit onderzoek slechts uit een klein aantal werknemers bestond dat zelf aangaf dat hun werkfunctioneren in belangrijke mate was veranderd, zijn de berekeningen in deze studie gebaseerd op kleine aantallen. De resultaten dienen daarom met

enige behoedzaamheid te worden beschouwd. Vervolgonderzoek om meer inzicht te krijgen in de interpreteerbaarheid van veranderingen in individuen daarom aan te raden. Voorlopig kunnen de berekende MIC waarden met voorzicht worden gebruikt.

Is een psychische module voor preventief medisch onderzoek voor verpleegkundigen en paramedici effectief in het stimuleren van hulpzoekgedrag en in het verbeteren van het functioneren in het werk en van de psychische gezondheid? (Hoofdstukken 6 en 7)

Een volgende stap in het onderzoek was de ontwikkeling van een psychische module preventief medisch onderzoek (PMO) voor verpleegkundigen en paramedici. Met de psychische module PMO werden werknemers gescreend op problemen in het werkfunctioneren en op aanwijzingen voor verschillende psychische klachten: stress, werkgerelateerde vermoeidheid, alcoholgebruik, depressie, angst en posttraumatische stressklachten. De persoonlijke resultaten van de screening werden direct online aan de werknemer teruggekoppeld. Werknemers die op één of meer aspecten positief werden gescreend, ontvingen een uitnodiging voor een preventief consult met hun bedrijfsarts. Hiervoor werd een 7-stappen protocol opgesteld, dat de bedrijfsartsen konden volgen. In de Appendix van dit proefschrift is een voorbeeld opgenomen voor de terugkoppeling van screeningsresultaten zoals werknemers deze hebben ontvangen. Ook is in de Appendix het protocol voor het preventieve consult van de bedrijfsartsen opgenomen.

In een cluster gerandomiseerde gecontroleerde studie werd de effectiviteit van de psychische module PMO getest ten opzichte van een controlegroep. De afdelingen van het universitair medisch centrum waar de studie werd uitgevoerd, werden willekeurig verdeeld over de PMO-groep (waarin de werknemers de psychische module PMO kregen aangeboden) en de controlegroep (waarin werknemers geen screeningsresultaten of bijzondere zorg kregen). Bij de start van het onderzoek, na drie maanden en na zes maanden werden de uitkomsten gemeten aan de hand van een online enquête.

In totaal hebben 191 verpleegkundigen en paramedici aan de psychische module PMO meegedaan. Hiervan werd 79% positief gescreend op één of meer problemen in het werkfunctioneren of psychische klachten. In de controlegroep, die uit 188 werknemers bestond, was dit 86%.

Hulpzoekgedrag is het bezoeken van één of meer hulpverleners in de afgelopen drie maanden. Hiervoor werd naar bezoek aan elf hulpverleners werd gevraagd, waaronder de huisarts, bedrijfsarts, fysiotherapeut, leidinggevende, psychologische hulpverleners, sociaal werkers en alternatieve hulpverleners. Het beloop in hulpzoekgedrag over de tijdsperiode van zes maanden verschilde statistisch significant tussen de twee groepen. Na drie maanden hadden in de PMO-groep 13% meer werknemers hulp gezocht dan in de controlegroep. Dit verschil berust voornamelijk op een stijging van bezoeken bij de bedrijfsarts, maar niet bij andere hulpverleners. Na zes maanden was het percentage van werknemers met

hulpzoekgedrag in de PMO-groep gedaald naar een niveau dat 8% lager was dan bij de controlegroep.

De PMO-groep vertoonde bij de vervolgmetingen minder problemen in werkfunctioneren dan de controlegroep. Van de diverse psychische klachten had het PMO alleen een effect op het beloop van alcoholgebruik over de tijd: tussen de drie en zes maanden na het begin van het onderzoek vertoonde de PMO-groep een daling in alcoholgebruik vergeleken met de controlegroep.

Er wordt aanbevolen om de screeningstrategie van de psychische module PMO aan te passen, omdat nu een hoog percentage werknemers wordt geselecteerd voor de uitnodiging voor een gesprek met de bedrijfsarts. Hiernaast wordt geadviseerd om de begeleiding door bedrijfsartsen te intensiveren met het doel meer effecten op psychische klachten te kunnen bereiken.

Hoe wordt de psychische module Preventief Medisch Onderzoek geëvalueerd door de gebruikers? (Hoofdstuk 8)

In een procesevaluatie van de psychische module PMO werd duidelijk dat 32% van de uitgenodigde werknemers deelnam aan de studie. Van de 151 verpleegkundigen en paramedici die positief gescreend waren, zijn 51 werknemers naar een preventief consult met hun bedrijfsarts geweest. Door een fout in het systeem werden 125 in plaats van 151 uitnodigingen voor de PMO-consulten verstuurd, er zijn dus 41% van de werknemers daadwerkelijk op deze uitnodiging ingegaan. Van deze groep geeft 80% aan dat de adviezen van de bedrijfsarts ter verbetering van het werkfunctioneren en de psychische gezondheid ook opgevolgd zijn. De bedrijfsartsen gaven aan dat 70% van de cliënten die zij ontvingen zich (gedeeltelijk) kon vinden in de screeningsresultaten over functioneren in het werk en 60% in de screeningresultaten over psychische klachten. Deelnemers waren tevreden over het consult met hun bedrijfsarts. Bijna iedereen had het gevoel open en eerlijk te kunnen zijn in het gesprek. De meerderheid van de werknemers zou het aanbod voor een psychische module PMO in de toekomst dan ook waarderen.

De bedrijfsartsen waren tevreden over het protocol voor de preventieve consulten en over de vooraf ontvangen training voor de psychische module PMO. Zij hebben het protocol voor het preventieve consult goed gevolgd; in de meeste PMO-consulten zijn alle zeven stappen doorlopen. De deelnemende bedrijfsartsen waren terughoudend in het verstrekken van adviezen. De bedrijfsartsen bleken het moeilijk te vinden om adviezen ter verbetering van klachten te geven aan werknemers die zelf geen duidelijke vraag om hulp hadden, maar puur op basis van de uitnodiging naar het consult zijn gekomen. De bedrijfsartsen hadden wel het gevoel dat de consulten een toegevoegde waarde hadden en dat de implementatie van de psychische module PMO in de toekomst waardevol is. Verbeterideeën van de bedrijfsartsen richten zich met name op de screeningstrategie, waarmee tot nu toe ook werknemers die slechts hele lichte klachten hebben, uitgenodigd werden voor het consult. Ook zijn

verbeteringen in de communicatiestrategie gewenst, waarbij de bedrijfsarts dit zelf regelt door contact te leggen met de afdelingen waarvoor hij/zij verantwoordelijk is.

Discussie en aanbevelingen

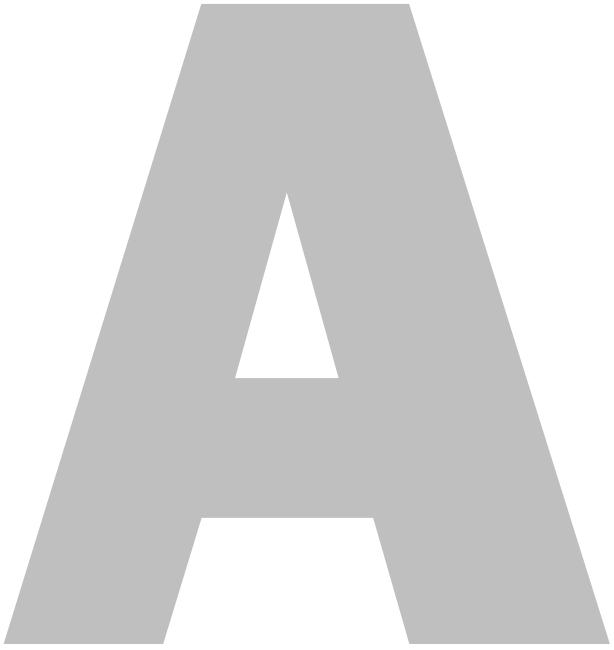
Op basis van de resultaten van dit proefschrift kan worden geconcludeerd dat de ontwikkeling van een beroepsspecifieke vragenlijst voor werkfunctioneren succesvol was. Zes van de zeven subschalen van de Nurses Work Functioning Questionnaire hebben een goede construcvaliditeit en reproduceerbaarheid voor het meten van verminderd werkfunctioneren onder verpleegkundigen en paramedici ten gevolge van psychische klachten. Vergeleken met bestaande, niet beroepsspecifieke vragenlijsten die werkfunctioneren meten, biedt de NWFQ belangrijke voordelen. Ten eerste omvat de NWFQ aspecten van werkfunctioneren die in andere vragenlijsten ontbreken, maar die wel belangrijk zijn bij het werken in de zorgsector, vanwege hun cruciale rol voor de patiëntenzorg en -veiligheid. Hieronder vallen het veroorzaken van fouten en het contact met patiënten en hun familie. Ten tweede geven de vragen in de NWFQ concrete voorbeelden van gedrag en werktaken, waardoor het invullen van de vragenlijst minder interpretatie van de invuller vereist. De concrete voorbeelden kunnen de reflectie van de werknemers op het eigen functioneren bevorderen. Er wordt verwacht dat dit de accuraatheid van de vragenlijst verbetert.

Er wordt geconcludeerd dat de NWFQ goed bruikbaar is in een psychische module PMO voor verpleegkundigen en paramedici, en dat de NWFQ nuttig is voor onderzoeksdoeleinden. Vervolgonderzoek zou zich moeten richten op verbetering van subschaal 2), *Moeite met besluitvorming,* en op het evalueren van afkappunten om werknemers met verminderd werkfunctioneren door psychische klachten beter te kunnen onderscheiden van werknemers zonder problemen.

Zoals beoogd, blijkt de psychische module PMO het hulpzoekgedrag te verhogen van verpleegkundigen en paramedici met indicaties voor verminderd werkfunctioneren of psychische klachten. Geconcludeerd wordt dat de ontwikkelde psychische module PMO de doelen van een PMO, zoals geformuleerd door de Nederlandse Vereniging Arbeids-en Bedrijfsgeneeskunde, gedeeltelijk bereikt. Het doel 'bewaken en bevorderen van het functioneren in het werk' is bereikt, aangezien werknemers die de psychische module PMO aangeboden kregen, een vermindering van problemen in het werkfunctioneren vertoonden. Het doel 'bewaken en bevorderen van de gezondheid' is onvoldoende bereikt. De psychische module PMO had alleen invloed op alcoholgebruik als psychische klacht. Overige psychische klachten, zoals stress, werkgerelateerde vermoeidheid, depressie, angst en posttraumatische stressklachten, namen niet af ten opzichte van de controlegroep.

Vervolgonderzoek kan zich richten op verbetering van de opzet van de screening en op verbetering van de daadwerkelijk geadviseerde interventies die de bedrijfsartsen aan de werknemers geven, zodat de psychische gezondheid verbeterd kan worden. Beleidsmakers in

universitair medische centra kunnen zich richten op het welzijn en de veiligheid van hun werknemers en patiënten, door het inzetten van een psychische module PMO. Voor de bedrijfsgezondheidszorg wordt aanbevolen om psychische modules PMO verder te ontwikkelen, te evalueren en om de implementatie hiervan te stimuleren, zowel binnen universitair medische centra als daarbuiten.



Appendix

APPENDIX 1

English version of the Nurses Work Functioning Questionnaire (NWFQ).

English version of the Nurses Work functioning Questionnaire

To what extent did you experience difficulty in meeting the following aspects **of your work** in the **past 4 weeks**?	**no difficulty**						**great difficulty**
1 Keeping sufficient overview of your tasks.	☐	☐	☐	☐	☐	☐	☐
2 Not forgetting something one day.	☐	☐	☐	☐	☐	☐	☐
3 Working efficiently.	☐	☐	☐	☐	☐	☐	☐
4 Carrying out your work activities in general.	☐	☐	☐	☐	☐	☐	☐
5 Carrying out your work independently.	☐	☐	☐	☐	☐	☐	☐
6 Staying alert in your work.	☐	☐	☐	☐	☐	☐	☐
7 Working accurately.	☐	☐	☐	☐	☐	☐	☐
8 Preventing incidents in your work.	☐	☐	☐	☐	☐	☐	☐
9 Working carefully.	☐	☐	☐	☐	☐	☐	☐
10 Contact with patients and their family.	☐	☐	☐	☐	☐	☐	☐
11 Showing sufficient empathy towards patients and their family.	☐	☐	☐	☐	☐	☐	☐
12 Taking time for your patients.	☐	☐	☐	☐	☐	☐	☐
13 Competently handling aggressive patients or family members.	☐	☐	☐	☐	☐	☐	☐
14 Error-free administering/handing medication.	☐	☐	☐	☐	☐	☐	☐

Relate the following statement **to your work** in the **past 4 weeks**. How much do you agree with it?	**totally disagree**	**disagree**	**disagree a little**	**not agree/ not disagree**	**agree a little**	**agree**	**totally agree**
15 I make mistakes more often than before.	☐	☐	☐	☐	☐	☐	☐

With respect to the **past 4 weeks**, are you someone who…	**(almost) never**	**once in a while**	**some-times**	**regularly**	**often**	**very often**	**(almost) always**
16 … has the feeling to have lost control over the work?	☐	☐	☐	☐	☐	☐	☐
17 … does not look forward to a working day/ shift?	☐	☐	☐	☐	☐	☐	☐
18 … starts the working day/ shift already moaning and groaning?	☐	☐	☐	☐	☐	☐	☐
19 … works without any enthusiasm?	☐	☐	☐	☐	☐	☐	☐
How often did you think the following in the **past 4 weeks**?							
20 I have great difficulty in getting through a working day.	☐	☐	☐	☐	☐	☐	☐
21 I have the need for an extra day off to be able to get through the working week well.	☐	☐	☐	☐	☐	☐	☐
How often did the following situations occur **in your work** in the **past 4 weeks**?							
22 I do not succeed in listening well to my patients.	☐	☐	☐	☐	☐	☐	☐
23 I notice myself that I treat patients too roughly.	☐	☐	☐	☐	☐	☐	☐

English version of the Nurses Work functioning Questionnaire (continued)

How often were the following situations **with respect to your work** applicable to you in the **past 4 weeks**?		**not once**	**1 x per month**	**2 - 3 x per month**	**on average 1 x per week**	**on average 2 - 3 x per week**	**on average 1 x per day**	**on average more than 1 x per day**
24	The tone of voice I use with my patients or their family is not too friendly.	□	□	□	□	□	□	□
25	I am curt towards patients or their family.	□	□	□	□	□	□	□
Answer the questions below with respect to the **past 4 weeks. How often did you...**								
26	... <u>almost</u> cause incidents in your work?	□	□	□	□	□	□	□
27	... <u>almost</u> make a mistake in the administration/handing of medicines?	□	□	□	□	□	□	□
28	... underestimated the seriousness of a situation?	□	□	□	□	□	□	□
How often did something go wrong while you were carrying out the tasks below in the **past 4 weeks**?								
29	Initiating infusion.	□	□	□	□	□	□	□
30	Assessing which (nursing) care a patient needs.	□	□	□	□	□	□	□
31	Performing and reporting actions (e.g. in the nursing plan or treatment plan).	□	□	□	□	□	□	□

How often did you **<u>almost</u>** do something wrong while carrying out the tasks below in the **past 4 weeks**?		**not once**	**1 x per month**	**2 - 3 x per month**	**on average 1 x per week**	**on average 2 - 3 x per week**	**on average 1 x per day**	**on average more than 1 x per day**
32	Performing and reporting actions (e.g. in the nursing plan or treatment plan).	□	□	□	□	□	□	□

Indicate how often in the **past 4 weeks** the following has happened.		**(almost) never**	**some-times**	**regularly**	**often**	**(almost) always**
33	Friction between you and someone from your team.	□	□	□	□	□
34	Being irritated during work.	□	□	□	□	□
35	Tensions between you and your colleagues.	□	□	□	□	□
How often was the behavior below applicable to you in the **past 4 weeks**?						
36	Leaving for a moment so you can be alone.	□	□	□	□	□
37	Avoiding conversations with your colleagues.	□	□	□	□	□
38	Avoiding common areas.	□	□	□	□	□
39	Avoiding working together with your colleagues.	□	□	□	□	□

English version of the Nurses Work functioning Questionnaire (continued)

Relate the following statements **to your work** in the **past 4 weeks**. How much do you agree with it?		**disagree**	**disagree a little**	**not agree/ not disagree**	**agree a little**	**agree**
40	Making decisions I leave to my colleagues.	□	□	□	□	□
41	I often only do what is absolutely necessary.	□	□	□	□	□
42	Meetings and evaluations I avoid as much as possible.	□	□	□	□	□
43	I prefer to do only routine jobs.	□	□	□	□	□
44	I am often astonished at how easy it is for others to upset me.	□	□	□	□	□
45	I often react irritated towards colleagues/managers.	□	□	□	□	□
46	I often get angry about matters at work.	□	□	□	□	□
47	I often have conflicts with my manager.	□	□	□	□	□
48*	I was able to make important decisions in my work responsibly. (R)	□	□	□	□	□
49*	I can quickly decide how to handle matters in my work. (R)	□	□	□	□	□
50*	I know how to make the right decisions in stressful situations. (R)	□	□	□	□	□

* Use of the items of subscale 2) *Impaired decision making*, is not recommended, as in their current form they have shown insufficient reproducibility.

R Items followed by (R) need to be recoded before the sum score is calculated.

Instructions for sum score calculation and categorization of sum scores

	Subscales	Items	Calculation of standardized sum score	# of items		Cut-off values for case classification		
				Total	Minimum # of items completed for sum score calculation	Green	Orange	Red
1	Cognitive aspects of task execution and general incidents	1, 2, 3, 4, 5, 6, 7, 8, 9, 15, 16	(sum of item scores x 100) / (# of items x 6)	11	9	0-21	22-25	25-100
2	Impaired decision making*	48(R), 49(R), 50(R)	(sum of item scores x 100) / (# of items x 4)	3	3	0-24	25-33	34-100
3	Causing incidents at work**	14, 26, 27, 28, 29, 30, 31, 32	(sum of item scores x 100) / (# of items x 6)	8	6	0-12	13-14	15-100
4	Avoidance behavior	36, 37, 38, 39, 40, 41, 42, 43	(sum of item scores x 100) / (# of items x 4)	8	6	0-9	10-12	13-100
5	Conflicts and irritations with colleagues	33, 34, 35, 44, 45, 46, 47	(sum of item scores x 100) / (# of items x 4)	7	6	0-17	18-28	29-100
6	Impaired contact with patients and their family	10, 11, 12, 13, 22, 23, 24, 25	(sum of item scores x 100) / (# of items x 6)	8	6	0-14	15-18	19-100
7	Lack of energy and motivation	17, 18, 19, 20, 21	(sum of item scores x 100) / (# of items x 6)	5	4	0-26	27-31	32-100
	Total score (without subscale 2)***	1 - 47	(sum of item scores x 100) / 252	47	┼	0-19	20-21	22-100

Technical details

- – Item score counting starts with 0 on the outer left category, add 1 point for each category further to the right (e.g., *disagree* = 0; *disagree a little* = 1; *not agree/not disagree* = 2; *agree a little* = 3; *agree* = 4).
- – R: items followed by (R) need to be recoded before the sum score is calculated.
- – Standardized sum scores are calculated according to the following principle: (sum of item scores x 100) / (# of items x maximum score per item).
- – For sum score calculation, subjects need to have filled out at least ¾ of all items of a subscale.
- – The range of the standardized sum score is 0-100 for each subscale.
- * Use of the items of subscale 2) *Impaired decision making*, is not recommended, as in their current form they have shown insufficient reproducibility.
- ** Subscale 3) *Causing incidents at work*, is not applicable to allied health professionals.
- *** For allied health professionals, the total NWFQ sum score can be calculated without the items of subscale 3 (thus with 39 items), according to the following principle: (sum of item scores x 100) / 204.
- ┼ At least ¾ of the items should be completed. In case of missing items the denominator of the standardized sum score calculation has to be adapted. In that case, the maximum sum score for the items competed has to be set in the denominator.

APPENDIX 2

Nederlandse versie van de Nurses Work Functioning Questionnaire (NWFQ). [Dutch version of the NWFQ]

Nederlandse versie van de Nurses Work Functioning Questionnaire

Hoeveel moeite hebben de volgende aspecten u **in uw werk** gekost in de **afgelopen 4 weken**?		**geen enkele moeite**						**heel veel moeite**
1	Voldoende overzicht houden over uw taken.	☐	☐	☐	☐	☐	☐	☐
2	Op een dag niets vergeten.	☐	☐	☐	☐	☐	☐	☐
3	Efficiënt werken.	☐	☐	☐	☐	☐	☐	☐
4	Uitvoering van uw werkzaamheden in het algemeen.	☐	☐	☐	☐	☐	☐	☐
5	Zelfstandig uitvoeren van uw werk.	☐	☐	☐	☐	☐	☐	☐
6	Alert blijven in uw werk.	☐	☐	☐	☐	☐	☐	☐
7	Nauwkeurig werken.	☐	☐	☐	☐	☐	☐	☐
8	Incidenten in uw werk voorkomen.	☐	☐	☐	☐	☐	☐	☐
9	Zorgvuldig werken.	☐	☐	☐	☐	☐	☐	☐
10	Contact met patiënten en hun familie.	☐	☐	☐	☐	☐	☐	☐
11	Tonen van voldoende empathie voor patiënten en hun familie.	☐	☐	☐	☐	☐	☐	☐
12	Tijd nemen voor uw patiënten.	☐	☐	☐	☐	☐	☐	☐
13	Goed omgaan met agressieve patiënten of familieleden.	☐	☐	☐	☐	☐	☐	☐
14	Foutloos toedienen/ aangeven van medicatie.	☐	☐	☐	☐	☐	☐	☐

Betrek de volgende uitspraak op uw werk in de **afgelopen 4 weken**. In hoeverre bent u het ermee eens?		**helemaal oneens**	**oneens**	**beetje oneens**	**niet eens/ niet oneens**	**beetje eens**	**eens**	**helemaal eens**
15	Ik maak vaker fouten dan voorheen.	☐	☐	☐	☐	☐	☐	☐

Met betrekking tot de **afgelopen 4 weken**, bent u iemand die...		**(bijna) nooit**	**een enkele keer**	**soms**	**regelmatig**	**vaak**	**heel vaak**	**(bijna) altijd**
16	... het gevoel heeft de controle over het werk kwijt te zijn?	☐	☐	☐	☐	☐	☐	☐
17	... opziet tegen een werkdag/ dienst?	☐	☐	☐	☐	☐	☐	☐
18	... al steunend en kreunend begint aan de werkdag/ dienst?	☐	☐	☐	☐	☐	☐	☐
19	... zonder enig enthousiasme werkt?	☐	☐	☐	☐	☐	☐	☐

Hoe vaak kwamen de volgende situaties **in uw werk** voor in de **afgelopen 4 weken**?								
20	Het kost me heel veel moeite om een werkdag door te komen.	☐	☐	☐	☐	☐	☐	☐
21	Ik heb de behoefte aan een extra vrije dag om de werkweek goed door te kunnen komen.	☐	☐	☐	☐	☐	☐	☐

Nederlandse versie van de Nurses Work Functioning Questionnaire (vervolgd)

Hoe vaak kwamen de volgende situaties **in uw werk** voor in de **afgelopen 4 weken**?		**(bijna) nooit**	**een enkele keer**	**soms**	**regelmatig**	**vaak**	**heel vaak**	**(bijna) altijd**
22	Het lukt me niet om goed naar mijn patiënten te luisteren.	□	□	□	□	□	□	□
23	Ik merk bij mezelf dat ik te ruw omga met patiënten.	□	□	□	□	□	□	□
Hoe vaak waren de volgende situaties **in het werk** in de **afgelopen 4 weken** op u van toepassing?		**geen enkele keer**	**1 x per maand**	**2 - 3 x per maand**	**gemiddeld 1 x per week**	**gemiddeld 2 - 3 x per week**	**gemiddeld 1 x per dag**	**gemiddeld meer dan 1 x per dag**
24	Mijn toon tegenover patiënten of hun familie is niet zo vriendelijk.	□	□	□	□	□	□	□
25	Ik ben kortaf naar patiënten of hun familie.	□	□	□	□	□	□	□
Beantwoord onderstaande vragen met betrekking tot de **afgelopen 4 weken. Hoe vaak heeft u....**								
26	... bijna incidenten in uw werk veroorzaakt?	□	□	□	□	□	□	□
27	... bijna een vergissing gemaakt bij het toedienen/ aanmaken van medicijnen?	□	□	□	□	□	□	□
28	... de ernst van een situatie verkeerd in geschat?	□	□	□	□	□	□	□
Hoe vaak ging er bij u iets mis bij de onderstaande taken in de **afgelopen 4 weken**?		**geen enkele keer**	**1 x per maand**	**2 - 3 x per maand**	**gemiddeld 1 x per week**	**gemiddeld 2 - 3 x per week**	**gemiddeld 1 x per dag**	**gemiddeld meer dan 1 x per dag**
29	Infuus instellen.	□	□	□	□	□	□	□
30	Inschatten van welke verpleegkundige zorg een patiënt nodig heeft.	□	□	□	□	□	□	□
31	Realisatie en documentatie van handelingen (b.v. in het verpleegkundig plan of behandelplan).	□	□	□	□	□	□	□
Hoe vaak ging er bij u **bijna** iets mis bij de onderstaande taken in de **afgelopen 4 weken**?								
32	Realisatie en documentatie van handelingen (b.v. in het verpleegkundig plan of behandelplan).	□	□	□	□	□	□	□
Geef aan hoe vaak in de **afgelopen 4 weken** het volgende is gebeurd.		**(bijna) nooit**	**soms**	**regelmatig**	**vaak**	**(bijna) altijd**		
33	Wrijving tussen u en iemand uit uw team.	□	□	□	□	□		
34	Geïrriteerd zijn tijdens het werk.	□	□	□	□	□		
35	Spanningen tussen u en uw collega's.	□	□	□	□	□		

Nederlandse versie van de Nurses Work Functioning Questionnaire (vervolgd)

Hoe vaak was het onderstaande gedrag op u van toepassing in de **afgelopen 4 weken**?		**(bijna) nooit**	**soms**	**regelmatig**	**vaak**	**(bijna) altijd**
36	Even weggaan zodat u alleen kunt zijn.	□	□	□	□	□
37	Gesprekken met uw collega's uit de weg gaan.	□	□	□	□	□
38	Gezamenlijke ruimtes vermijden.	□	□	□	□	□
39	Samenwerking met uw collega's vermijden.	□	□	□	□	□

Betrek de volgende uitspraken **op uw werk** in **de afgelopen 4 weken.** In hoeverre bent u het ermee eens?		**oneens**	**beetje oneens**	**niet eens/ niet oneens**	**beetje eens**	**eens**
40	Besluiten nemen laat ik aan collega's over.	□	□	□	□	□
41	Ik doe vaak alleen het hoognodige.	□	□	□	□	□
42	Vergaderingen en evaluaties vermijd ik zo veel mogelijk.	□	□	□	□	□
43	Het liefst doe ik alleen routine klussen.	□	□	□	□	□
44	Ik verbaas me er vaak over hoe weinig ik van anderen kan hebben.	□	□	□	□	□
45	Ik reageer vaak geïrriteerd op collega's/ leidinggevende.	□	□	□	□	□
46	Ik word vaak boos over zaken op het werk.	□	□	□	□	□
47	Ik heb vaak conflicten met mijn leidinggevende.	□	□	□	□	□
48*	Belangrijke besluiten in het werk kan ik goed zelf nemen. (R)	□	□	□	□	□
49*	Ik kan snel beslissen hoe ik zaken in mijn werk moet aanpakken. (R)	□	□	□	□	□
50*	In stressvolle situaties weet ik de juiste besluiten te nemen. (R)	□	□	□	□	□

* Het wordt afgeraden om de items van subschaal 2) *Moeite met besluitvorming*, in hun huidige vorm te gebruiken, vanwege een gebrek aan betrouwbaarheid van deze subschaal.

R Items met een (R) dienen gespiegeld te worden voor het berekenen van de somscores.

Instructies voor het berekenen van de somscores en presentatie van de afkappunten.

Subschaal		Items	Berekening van de gestandaardiseerde som score	# items		Afkappunten		
				Totaal	Minimum # ingevulde items voor som score berekening	Groen	Oranje	Rood
1	Cognitieve aspecten van taak uitvoering en algemene fouten	1, 2, 3, 4, 5, 6, 7, 8, 9, 15, 16	(som van item scores x 100) / (# items x 6)	11	9	0-21	22-25	25-100
2	Moeite met besluitvorming*	48(R), 49(R), 50(R)	(som van item scores x 100) / (# items x 4)	3	3	0-24	25-33	34-100
3	Veroorzaken van incidenten in het werk**	14, 26, 27, 28, 29, 30, 31, 32	(som van item scores x 100) / (# items x 6)	8	6	0-12	13-14	15-100
4	Vermijdingsgedrag	36, 37, 38, 39, 40, 41, 42, 43	(som van item scores x 100) / (# items x 4)	8	6	0-9	10-12	13-100
5	Conflicten en irritatie met collega's	33, 34, 35, 44, 45, 46, 47	(som van item scores x 100) / (# items x 4)	7	6	0-17	18-28	29-100
6	Verminderde omgang met patiënten en hun familie	10, 11, 12, 13, 22, 23, 24, 25	(som van item scores x 100) / (# items x 6)	8	6	0-14	15-18	19-100
7	Gebrek aan energie en motivatie	17, 18, 19, 20, 21	(som van item scores x 100) / (# items x 6)	5	4	0-26	27-31	32-100
	Totaal score (zonder subschaal 2)***	1 - 47	(som van item scores x 100) / 252	47	┼	0-19	20-21	22-100

Technische details

- Item scores beginnen met 0 in de meest linkse antwoord categorie, voor elke categorie verder naar rechts wordt 1 punt opgeteld (bv. *oneens* = 0; *beetje oneens* = 1; *niet eens/niet oneens* = 2; *beetje eens* = 3; *eens* = 4).
- R: Items met een (R) dienen gespiegeld te worden voor het berekenen van de somscores.
- Berekening van de gestandaardiseerde somscores gaat volgens het principe: (som van de itemscores x 100) / (# items x maximale itemscore).
- Voor de berekening van som scores dient minimaal ¾ van de items van de subschaal te zijn ingevuld.
- De range van de gestandaardiseerde somscores is 0-100 voor elke subschaal.

* Het wordt afgeraden om de items van subschaal 2) *Moeite met besluitvorming*, in hun huidige vorm te gebruiken, vanwege een gebrek aan betrouwbaarheid van deze subschaal.

** Subschaal 3) *Veroorzaken van incidenten in het werk*, is alleen geschikt voor verpleegkundigen en operatiemedewerkers, niet voor paramedici.

*** Voor paramedici kan een totaal somscore worden berekend zonder de items van subschaal 3 (dus met 39 items). De berekening gaat volgens het principe: (som van item scores x 100) / 204.

┼ Minimaal ¾ van de items dient te zijn ingevuld. Bij missende items dient de noemer van de berekening voor de gestandaardiseerde somscores te worden aangepast. In dit geval wordt de maximale somscore van de ingevulde items als noemer ingevuld.

APPENDIX 3

Overzicht van de psychische module PMO [Overview of the WHS mental module]

Stap 1: De screening

- Deelnemers worden online gescreend op psychische klachten en verminderd werkfunctioneren (zie Appendix 3a).
- Deelnemers krijgen online een terugkoppeling van de persoonlijke screeningsresultaten (zie Appendix 3b).

Stap 2: Het PMO-consult met de bedrijfsarts

- Werknemers die positief zijn gescreend ontvangen een schriftelijke uitnodiging thuis voor het consult met hun bedrijfsarts (zie Appendix 3c).
- Het PMO-consult met de bedrijfsarts duurt 30 minuten en wordt gevoerd aan de hand van een 7-stappen protocol. In het protocol voor de bedrijfsartsen zijn de uitgangspunten voor het PMO-consult beschreven en wordt een overzicht gegeven van de 7-stappen van het consult en een overzicht van mogelijke interventies (zie Appendix 3d).

APPENDIX 3a

Overzicht van de screeningsinstrumenten

Zes type psychische klachten

1. Stress (4DKL-stress)
2. Werkgerelateerde vermoeidheid (VBBA Herstelbehoefte)
3. Alcoholmisbruik (AUDIT-C)
4. Depressie (BSI-DEP)
   4.1 Suïciderisico
5. Angst (BSI-ANG)
   5.1 Paniek (PHQ-15)
6. Posttraumatische stressstoornis (SVL)

Zeven aspecten van verminderd werk functioneren, d.m.v. de NWFQ

1. Cognitieve aspecten van taak uitvoering en algemene fouten
2. Moeite met besluitvorming*
3. Veroorzaken van incidenten in het werk**
4. Vermijdingsgedrag
5. Conflicten en irritatie met collega's
6. Verminderde omgang met patiënten en hun familie
7. Gebrek aan energie en motivatie

* Gebruik van deze subschaal wordt voor de toekomst afgeraden.

** Deze subschaal is alleen van toepassing op verpleegkundigen, operatieassistenten en anesthesiemedewerkers, niet op paramedici en verpleegkundig consulenten.

APPENDIX 3b

Voorbeeld terugkoppeling van screeningsresultaten voor medewerkers met, werkgerelateerde vermoeidheid en verminderd werk functioneren op een aantal aspecten

Geachte <<*aanhef en naam werknemer*>>,

U heeft zojuist de vragenlijst ingevuld van het onderzoek Mental Vitality @ Work. Hieronder geven wij u een overzicht van uw persoonlijke resultaten.

Eerst ziet u de resultaten over uw functioneren in uw werk. Hierna volgen de resultaten over uw mentale gezondheid. Aan het einde van dit bericht informeren wij u over mogelijkheden die er zijn om u verder te helpen, mocht dat nodig zijn.

FUNCTIONEREN IN HET WERK
De vragen die wij u hebben gesteld over uw functioneren in uw werk hebben betrekking op zeven onderdelen. Dit zijn onderdelen van het werk die beïnvloed kunnen worden door de mentale gezondheid van werknemers.

Uw persoonlijke resultaten komen tot stand door een vergelijking van uw antwoorden met de gemiddelde antwoorden van een grote groep verpleegkundigen, operatieassistenten, anesthesiemedewerkers en paramedici.

Bij elk onderdeel zijn er drie mogelijke resultaten: groen, oranje of rood.

Groen betekent dat uw functioneren op het betreffende onderdeel goed is. Op basis van uw antwoorden lijkt er geen reden te zijn om iets te veranderen.

Bij oranje lijkt het erop dat u op dit moment op het betreffende onderdeel iets minder goed scoort dan uw collega's. Mogelijk is bij dit onderdeel verbetering mogelijk.

Rood wil zeggen dat het erop lijkt dat het betreffende onderdeel van het werk u op dit moment minder goed af gaat. Uw scores zijn dus minder goed dan die van uw collega's. Dit onderdeel is voor u mogelijk een aandachtspunt.

Uw persoonlijke resultaten voor functioneren in uw werk
De vragenlijst over het functioneren in het werk bevat verschillende onderdelen. Hier volgen uw resultaten per aspect:

1) Cognitieve aspecten van taak uitvoering en algemene incidenten: oranje
Dit onderdeel omvat vragen over concentratie, overzicht houden, veroorzaken van algemene fouten en efficiënt, alert en zorgvuldig werken.

2) Moeite met besluiten nemen: groen
Dit onderdeel omvat vragen over het snel en goed kunnen nemen van besluiten in stressvolle situaties.

3) Veroorzaken van incidenten in het werk: oranje
Dit onderdeel omvat vragen over het veroorzaken of het bijna veroorzaken van incidenten in het werk, zoals bij het toedienen of aanmaken van medicijnen, instellen van infuus, inschatten van ernstige situaties, realisatie en documentatie van handelingen.

4) Vermijdingsgedrag vertonen: oranje
Dit onderdeel omvat vragen over het vermijden van taken of contact met collega's/leidinggevende.

5) Conflicten en irritaties met collega's: groen
Dit onderdeel omvat vragen over wrijvingen en irritaties met collega's/leidinggevende. Ook is er aandacht voor het boos worden over zaken op het werk.

6) Verminderde omgang met patiënten en hun familie: groen
Dit onderdeel omvat vragen over geduld en empathie voor patiënten en hun familie. Bijvoorbeeld goed luisteren naar patiënten en hun familie en de omgang met patiënten.

7) Gebrek aan energie en motivatie: rood
Dit onderdeel omvat vragen over moeite hebben om een werkdag of werkweek door te komen. Ook het opzien tegen werkdagen en zonder enthousiasme werken komen aan bod.

Uw antwoorden leiden op de onderdelen van functioneren in het werk tot enkele groene resultaten, enkele oranje resultaten en één rood resultaat. Het lijkt dus op een aantal aspecten goed te gaan met uw functioneren in het werk, maar op een aantal onderdelen op dit moment minder goed. Ga voor uzelf na of u zich kunt vinden in het resultaat oranje en rood voor de betreffende onderdelen. Misschien dat u een duidelijke oorzaak voor deze scores weet. U zou aan deze aspecten van uw werk de komende tijd extra aandacht kunnen besteden. Onderaan dit bericht komen wij hier op terug.

MENTALE GEZONDHEID

Uw persoonlijke resultaten voor uw mentale gezondheid

De vragen over uw mentale gezondheid gaan in op een aantal verschillende mentale gezondheidsklachten. De aspecten waarnaar is gekeken zijn stress, vermoeidheid na een dag werken, alcoholgebruik, somberheid, angstige gevoelens en klachten naar aanleiding van een ingrijpende gebeurtenis.

Uit uw antwoorden op de vragenlijst blijkt dat vermoeidheid na een dag werken mogelijk een aandachtspunt voor u is. Hieronder informeren wij u over mogelijkheden om uw mentale gezondheid te verbeteren.

Hoe verder?

Op basis van uw resultaten zijn uw mentale gezondheid en/of uw functioneren in uw werk mogelijk aandachtspunten voor u.

Het is bekend dat u een beroep heeft met een aanzienlijke mentale belasting. Daarnaast is optimaal functioneren in uw werk bijzonder belangrijk voor de gezondheid van uw patiënten, maar ook voor uw eigen gezondheid (denk bijvoorbeeld aan ongelukken in het werk zoals prikaccidenten). Aandacht voor uw mentale gezondheid en functioneren is daarom van extra belang om het ontstaan of het verergeren van mogelijke problemen te voorkomen.

Wanneer u uw functioneren in het werk en mentale gezondheid wilt verbeteren adviseren wij u om met anderen over uw resultaten uit dit onderzoek te praten en mogelijke oplossingen te bedenken. Dat kan bijvoorbeeld met uw leidinggevende, een coach of uw bedrijfsarts.

In het kader van het onderzoek Mental Vitality @ Work ontvangt u binnen 2 weken een uitnodiging voor een gesprek met uw eigen bedrijfsarts van de <<*naam uitvoerende arbodienst*>>. Deze kan mogelijke problemen met u bespreken en kan u adviezen geven. De uitnodiging is geheel vrijblijvend. Het onderzoeksteam stelt verder niemand (dus ook niet uw leidinggevende) van de uitnodiging of van de inhoud van het gesprek op de hoogte.

Mocht u zelf contact op willen nemen met de <<naam uitvoerende arbodienst>>, dan kan dit via:
E-mail: <<e-*mailadres uitvoerende arbodienst*>>
Telefoonnummer: <<*telefoonnr. uitvoerende arbodienst*>>

Met vriendelijke groet,

De medewerkers van de <<*uitvoerende arbodienst*>>

APPENDIX 3c

Uitnodigingsbrief voor het PMO-consult

<<aanhef en naam werknemer>>
<<adres werknemer>>
<<woonplaats werknemer>>

<<naam arbodienst>>
<<locatie arbodienst>>
datum : *<<datum>>* *<<telefoonnr. arbodienst>>*
uw kenmerk : onderzoek Mental Vitality @ Work *<<e-mail arbodienst>>*
ons kenmerk : *<<kenmerk>>* *<<faxnr. arbodienst>>*

betreft: uitnodiging spreekuur in het kader van Mental Vitality @ Work

Geachte *<<aanhef en naam werknemer>>*,
U neemt deel aan het onderzoek Mental Vitality @ Work, waarvoor hartelijk dank. In het kader van dit onderzoek nodigt uw bedrijfsarts, *<<aanhef en naam bedrijfsarts>>*, u uit voor het spreekuur:

op	*<<datum>>*
om	*<<tijdstip>>*
ruimte	*<<locatie>>*

Tijdens het spreekuur van een half uur gaat de bedrijfsarts de resultaten uit de vragenlijst van het onderzoek Mental Vitality @ Work met u doornemen. Er worden mogelijke problemen in uw mentale gezondheid of in uw werk besproken en uw bedrijfsarts kan adviezen hierover geven.

Deze uitnodiging en het spreekuur worden geheel vertrouwelijk behandeld. Het onderzoeksteam stelt verder niemand (dus ook niet uw leidinggevende) van deze uitnodiging of van de inhoud van het spreekuur op de hoogte.
De uitnodiging voor een gesprek met uw bedrijfsarts is geheel vrijblijvend. Indien u verhinderd bent of besluit niet met uw bedrijfsarts te willen praten, wordt u verzocht zo spoedig mogelijk telefonisch contact op te nemen met de *<<naam administratie van de uitvoerende arbodienst>>* voor het verplaatsen of het afzeggen van de afspraak, via telefoon *<<telefoonnummer arbodienst>>*.

Met vriendelijke groet,

<<naam medewerker>>
<<naam uitvoerende arbodienst>>

APPENDIX 3d

Protocol voor het consult bij de psychische module PMO in de zorg

Uitgangspunten

Doel van het consult

Werknemers waarbij op basis van diagnostische instrumenten in een PMO-vragenlijst blijkt dat er problemen zijn met hun psychische gezondheid en/of in hun functioneren in het werk worden uitgenodigd voor het PMO-consult. Doel is om bij deze groep duurzame inzetbaarheid, dus het goed en gezond werken, te bevorderen door tijdig hulp te bieden. Het consult, van 30 minuten, dient ertoe om het probleem van de werknemer en oorzaken hiervan te verhelderen, om perspectief te bieden en zicht te geven op mogelijke oplossing/interventies.

Rol van de bedrijfsarts

Het PMO-consult zal voor een aantal medewerkers de eerste keer zijn dat zij contact hebben met de bedrijfsarts (BA). Belangrijk is dat tijdens dit consult de werknemer kennis maakt met de BA als een vertrouwenspersoon en aanspreekpunt voor problemen rondom arbeid en gezondheid. U kunt als BA uw rol uitleggen en de vertrouwelijkheid van het contact benadrukken.

Het PMO-consult dient ertoe om het proces van hulp krijgen voor psychische klachten en verminderd werk functioneren in gang te zetten. De procesrol, met probleemverheldering, perspectief bieden en vervolgens het proces bewaken, staat dan ook op de voorgrond. De interventierol wordt flexibel ingevuld, afhankelijk van de behoefte van de werknemer en afhankelijk van de werkstijl en voorkeuren van de BA.

De adviezen

De BA heeft voor de begeleiding van werknemers met psychische klachten en hieruit voortkomende problemen in het werk functioneren verschillende adviezen en interventies ter beschikking. De keuze hieruit is afhankelijk van de specifieke symptomen en klachten, de ernst hiervan en de voorkeur van de werknemer zelf. Wat betreft adviezen gericht op de persoon zelf en de psychische klachten worden hier alleen algemene voorbeelden gegeven, omdat dit onderdeel nauw aansluit bij uw huidige praktijk. Voor mogelijke adviezen gericht op functioneren in het werk is een overzicht aan dit protocol toegevoegd. Per aspect van verminderd werk functioneren dat in de PMO-screening gemeten wordt, wordt in dit overzicht aangegeven welke mogelijke adviezen van toepassing zijn en hoe concrete voorbeelden eruit zien. De BA kan hier een keuze uit maken. Ook beslissingen om al dan niet door te verwijzen naar andere zorgprofessionals worden aan u zelf overgelaten.

Hanteren van het 7 stappen protocol

In het onderstaande zeven stappen protocol zijn alle belangrijke onderdelen van het PMO-consult weergegeven. Dit protocol dient als leidraad voor het (eerste) consult. Het protocol geeft zo veel mogelijk de manier van aanpak weer die nu bij open spreekuur consulten door u worden gehanteerd. De werknemers krijgen op basis van de resultaten in een screeningsvragenlijst een uitnodiging voor dit consult. Wij verwachten daarom dat zij mogelijk niet helder voor ogen hebben wat de BA voor hen kan betekenen. Ook zal de werknemer mogelijk een minder heldere hulpvraag hebben dan u gewend bent.

Voor het PMO-consult geldt dat de begeleiding van de werknemers 'maatwerk' is. De inhoudelijke invulling van het protocol dient flexibel te zijn afhankelijk van het probleem en de situatie van de individuele werknemer. Ook geldt dat de precieze uitvoering van de beschreven stappen afhankelijk is van uw persoonlijke werkstijl.

De 7 stappen van het PMO consult en hun invulling

Stap 1 Welkom & verwachtingen verhelderen

- Welkom heten.
- Uitleggen van doel PMO-consult en de eigen rol.
- Vertrouwelijk karakter van het consult benadrukken, BA noemt beroepsgeheim en legt uit dat leidinggevende alleen geïnformeerd wordt als hiervoor expliciet toestemming is gegeven.
- Duur en aard van dit consult uitleggen.
- Vraag naar verwachting van de WN voor dit consult. Eventuele verkeerde verwachtingen bijstellen.

Stap 2 Resultaten bespreken

- Resultaten uit de vragenlijst wat betreft psychische klachten en wat betreft werk functioneren met de werknemer bespreken.
- Check of WN de resultaten begrijpt.
- Check of WN zich in de resultaten kan vinden.
    - Zo nee, bespreken hoe deze resultaten tot stand zijn gekomen
        - Verken of er andere aanwijzingen zijn voor problemen in de mentale gezondheid of in het werk. Zoja, bevorder reflectie hierover bij WN en ga verder met de overige stappen van dit protocol.
        - Als er geen andere aanwijzingen zijn voor klachten en verminderd werk functioneren, leg dan in algemenere zin uit wat gedaan kan worden mochten in de toekomst problemen ontstaan en wat een bedrijfsarts hierbij kan betekenen.
    - Zo ja, bespreken
        - Hoe de klachten/verminderd werk functioneren zich uiten
        - De tijdslijn van de klachten/verminderd werk functioneren
        - De waargenomen ernst van de klachten/verminderd werk functioneren
        - Met wie er over de klachten/verminderd werk functioneren is gesproken
            - Collega's/leidinggevende
            - Partner, familie en vrienden
        - Wat WN zelf al heeft gedaan om klachten/verminderd werk functioneren te verminderen
        - of er contact is met andere zorgverleners, bv. huisarts of een psycholoog

Stap 3 Probleeminventarisatie

- Verken mogelijke oorzaken voor het ontstaan van de klachten. Bespreek hiervoor drie aspecten van de leef- en werksituatie van de WN.
  - Werksituatie
  - Privésituatie
  - Medische situatie
- Verken de gevolgen van aanwezige klachten voor het werk functioneren. Gebruik hierbij eventueel ook de resultaten van de PMO vragenlijst over werk functioneren.

Stap 4 Probleemdiagnose & rationale geven

- Vat de hulpvraag van de WN helder samen en verifieer of deze juist is.
  - Benoem de oorzaken voor klachten en verifieer deze.
- Geef voorlichting over de klachten en het ontstaan ervan. Bied de wn perspectief d.m.v. rationale.

Stap 5 Adviezen bespreken

- Geef afhankelijk van de hulpvraag adviezen.
  - Ga eerst na wat de WN zelf het meest nodig heeft en nuttig vindt en pas de adviezen hierop aan.
  - Adviezen kunnen zich richten op
    - De persoon (en zijn/haar klachten), de privé belasting
    - Het werk (inzetbaarheid en functioneren in het werk)
- Bespreek met WN de communicatie met zijn/haar leidinggevende over klachten en adviezen.
  - Stimuleer de WN om zelf contact met leidinggevende op te nemen over de klachten.
  - Geef mogelijke adviezen/tips voor het gesprek met de leidinggevende, bereid het gesprek zo nodig samen voor.

Stap 6 Vervolg

- Maak heldere afspraken met WN over het vervolg.
  - Verhelder welke rol u als BA kunt blijven houden in het vervolg. Maak eventueel een vervolgafspraak.
  - Bespreek zo nodig mogelijke doorverwijzingen met de WN, bv. naar de huisarts, (bedrijfs)maatschappelijk werker of psycholoog.

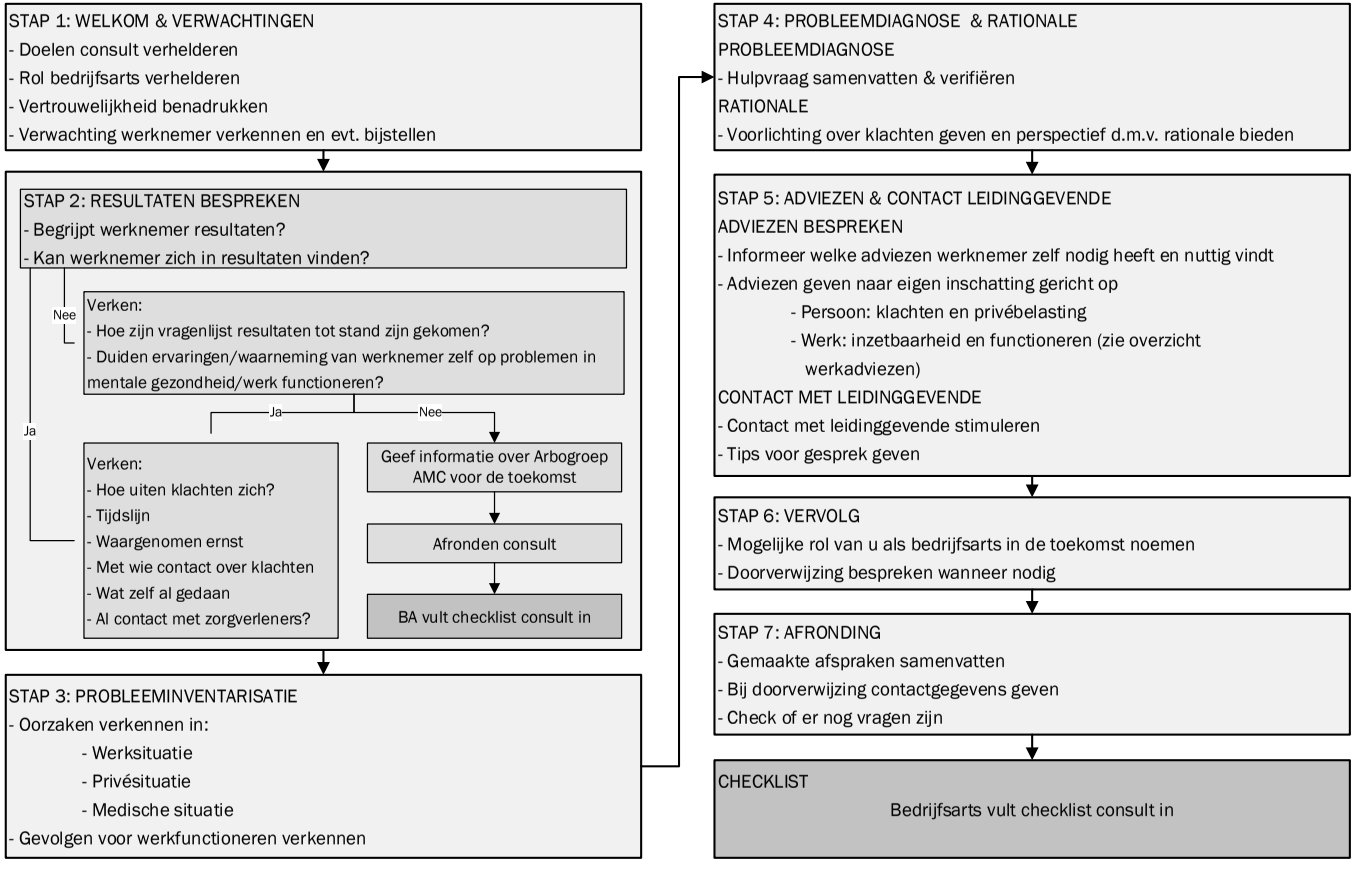
Stap 7 Afronding

- Vat afspraken die gemaakt zijn over contact met de leidinggevende samen.
- Bij doorverwijzing check of de WN weet met wie en hoe hij/zij contact moet opnemen.
- Check of de WN nog vragen heeft.

BA = bedrijfsarts
WN = werknemer

MENTAL VITALITY @ WORK

PSYCHISCH MODULE PMO IN DE ZORG – HET 7 STAPPEN CONSULT PROTOCOL



Overzicht van mogelijke interventies

Gericht op de persoon & de klachten

Wat betreft deze adviezen geven wij minder specifieke voorbeelden, omdat dit nauw aansluit bij uw huidige praktijk. Mogelijke adviezen gericht op de persoon en zijn/haar klachten kunnen zich op verschillende aspecten richten. Deze zijn:

1. Stress management: dagstructuur (actief ontspannen), slaapstructuur, piekeropdracht, taakanalyse
2. Coping: positief her-etiketteren, nee-zeggen, perspectief geven
3. Leefstijlinterventies: lichaamsbeweging, gezonder eetpatroon, minder alcoholgebruik
4. Belasting in de privésfeer reduceren.

Gericht op de gevolgen voor het werk functioneren:

Adviezen voor het verbeteren van werk functioneren richten zich in de eerste plaats op het tijdelijk verminderen van risico's in het werk bv. voor het veroorzaken van incidenten en op het verminderen van extra belasting, totdat de klachten in de psychische gezondheid zijn verbeterd.

Het is aan te bevelen om eerst samen met de werknemer aanwezige problemen zo concreet mogelijk te formuleren en te verkennen wat de werknemer zelf aan oplossingsmogelijkheden heeft. Neem hierin ook de vraag mee "of" een probleem daadwerkelijk opgelost kan worden en of dat op korte termijn kan gebeuren of dat het van latere orde is. Adviezen die u uiteindelijk kunt geven kunnen volgende maatregelen inhouden:

A. De werkinhoud aanpassen, bv. door specifieke taken in te perken.
B. Het werk anders organiseren, bv. tijdelijke aanpassing in de volgorde van taken; meer pauzes nemen; juist wel of juist niet samenwerking met collega's bevorderen; meer supervisie op specifieke taken krijgen, ondersteuning regelen bv. van collega's of ondersteunend personeel.
C. Het bijstellen van werktijden, bv. geen nachtdiensten werken.
D. Het verminderen van de hoeveelheid werkuren.

In de tabel op de volgende pagina vindt u een overzicht waarin geschikte regelingen en voorbeelden zijn weergegeven voor de zeven aspecten van verminderd werk functioneren waarop gescreend wordt.

Psychische module PMO in de zorg – Overzicht van mogelijke werkadviezen

	Aangrijpingspunten v/d interventie				Voorbeelden
	A) Werk-inhoud	B) Organi-satie v/h werk	C) Werk-tijden	D) Werk-uren	
1. Cognitieve aspecten van taak uitvoering en algemene fouten	+	+	+	+	A) Taken met veel kans op fouten niet vervullen. B) Complexe en simpele taken afwisselen, of complexe taken inplannen op momenten van de dag dat er minder cognitieve problemen zijn of als er minder verstoringen zijn. B) Meer rustmomenten, pauzes inlassen. B) Handelingen laten controleren door collega's. B) Voor bepaalde taken een rustige plek met weinig kans op onderbrekingen opzoeken. C & D) Minder uren of op minder belastende tijden werken.
2. Moeite met besluitvorming	+	+	-	-	A) Taken tijdelijk aanpassen, bv. minder complexe patiënten, meer routine taken. B) Meer tijd inbouwen voor taken, en meer tijd nemen voor het nemen van beslissingen. B) Meer overleg met andere collega's of de leidinggevende.
3. Veroorzaken van incidenten in het werk	+	+	-	-	A) Taken met veel kans op fouten tijdelijk niet uitvoeren, zoals medicijnen uitdelen. B) Meer overleg- en controlemomenten met collega's/leidinggevende inbouwen. B) Voor bepaalde taken een rustige plek met weinig kans op onderbrekingen opzoeken.
4. Vermijdingsgedrag	+	+	-	-	A) Bewust juist die taken doen die u vermijdt. Stel ter beloning hier iets leuks tegenover als u ze gedaan heeft. B) Ter motivatie vaker samenwerken met positief ingestelde collega's. B) Met jezelf contactmomenten afspreken bv. elke werkdag één keer een praatje maken met een collega of afspreken om samen met collega's te lunchen.
5. Conflicten en irritatie met collega's	+	+	-	-	A & B) Werk toespitsen op taken waarin samenwerking minder noodzakelijk is. B) Collega's om feedback vragen over hoe zij de omgang van de werknemer met anderen waarnemen en bewust bedenken waarop wn in de toekomst actiever op kan letten.
6. Verminderde omgang met patiënten en hun familie	+	+	-	-	A) Kiezen voor taken waarin minder patiëntencontact plaats vindt. B) Laten samenwerken met andere collega's bij taken met veel patiënten contact. B) Collega's om feedback vragen over hoe zij de omgang van de werknemer met patiënten waarnemen en bewust bedenken waarop wn in de toekomst actiever op kan letten.
7. Gebrek aan energie en motivatie	+	+	+	+	A & B) Een lijst van energievreters en energiegevers in het werk samenstellen. Meer energiegevende taken of activiteiten in bouwen in werk en aan het einde van de dag terug te denken wat de energiegevers van die dag waren. B) Ter motivatie veel laten samenwerken met positief ingestelde collega's. B) Pauzes inplannen waarin juist iets fysiek actiefs word gedaan, zoals een lunchwandeling buiten. C & D) Minder uren of op minder belastende tijden werken.

D

Dankwoord

DANKWOORD

Dit proefschrift was zonder de ondersteuning van velen niet tot stand gekomen. Graag wil ik alle betrokkenen bij het onderzoek hartelijk danken en de volgende personen in het bijzonder.

Beste promotor en copromotores, hartelijk dank voor jullie enorme betrokkenheid bij het onderzoeksproject en bij mij persoonlijk. **Frank**, jouw creatieve en kritische geest waren in de afgelopen vier jaar onmisbaar voor mij. **Karen**, je oneindige enthousiasme en bevlogenheid waren een onontbeerlijke bron van energie en inspiratie. **Judith**, je gedrevenheid, scherpe blik en je nooit aflatende vertrouwen tijdens de laatste maanden waren voor mij van essentieel belang.

Stichting Instituut GAK, hartelijk dank voor het vertrouwen in het belang van dit onderzoeksproject dat jullie met de subsidie hebben laten blijken.

Graag bedank ik de leden van de **promotiecommissie**, Prof. dr. J.R. Anema, Dr. U. Bültmann, Prof. dr. P.M.G Emmelkamp, Prof. dr. R.J. de Haan, Prof. dr. H.C.W. de Vet en Prof. dr. H.P.C.M. van Weert, voor het beoordelen van mijn manuscript en de bereidheid tijdens de aanstaande verdediging te opponeren.

Ook wil ik mijn dank uitspreken aan de leden van de **begeleidingscommissie** in het AMC, voor alle inhoudelijke en praktische ondersteuning: G. Frijstein, C.A Moleman, C.J. Bijl en F. Nollet, H.P.M.N. van Wersch en A.K. Roelofs.

Mijn dank gaat ook naar de **Arbogroep AMC**, beste **Gerard**, beste **bedrijfsartsen**, jullie medewerking was van cruciaal belang voor het slagen van mijn promotietraject. Hartelijk dank voor alle energie en tijd die jullie in het onderzoek hebben gestoken en voor de inhoudelijke input en feedback over het PMO.

Sarah, jij was mijn steun en toeverlaat tijdens het uitvoeren van de RCT. Ik had jouw oog voor detail, je tomeloze energie en je gezelligheid niet willen missen. Jij hebt de tweede helft van mijn promotietraject zo veel beter en leuker gemaakt. Dankjewel!

Beste **Trimbos collega's**, Eva, Iris, Jolanda, Linda en Odile, de samenwerking met jullie was een grote verrijking voor dit onderzoek. Hartelijk dankt voor de leuke en goede samenwerking.

Coronel collega's, jullie allen wil ik hartelijk danken voor jullie aandeel in de leuke tijd die ik in de laatste vier jaar heb gehad. Jullie deuren staan altijd open voor tips en inspiratie, maar ook voor de nodige afleiding! Een aantal collega's wil ik hierbij in het bijzonder noemen: Martijn, Maaike, Julitta, Vincent, Inge en Thea. Lieve **Sietske**, **Margot**, **Marrit**, **Sarah**, **Marie-Christine**, **Marieke**, **Iris**, en **Willemijn,** in de laatste jaren zijn jullie veel meer dan collega's geweest. De vriendschappen met jullie betekenen heel veel voor mij! Dank jullie wel!

Beste collega's van **Medische Besliskunde, LUMC**, jullie hebben mij in april open als nieuwe collega ontvangen. Dank jullie wel daarvoor! Ik geniet van de boeiende en leuke onderzoeksomgeving bij jullie en hoop er nog lang deel van uit te maken.

Lieve **Selma**, **Julia**, **Moritz**, **Karen**, **Hannah** en **Heleanne**, onze vriendschappen zijn onmisbaar voor mij, door jullie eindeloze betrokkenheid en ondersteuning en natuurlijk door samen te genieten van het leven. Weten dat ik altijd op jullie kan vertrouwen is van onschatbare waarde. Dank jullie wel!

Beste **Dennis**, jij was in de afgelopen jaren vaste gast aan de Platanenweg en Drossestraat en wij hebben wel en wee gedeeld boven biertjes, kippen en menig spelletje. Dank voor de goede vriendschap en natuurlijk voor je creatieve bijdrage aan dit boek!

Beste **Arnout**, jouw 24/7 service voor taal en ICT en al je goede zorgen in de afgelopen maanden waren erg waardevol voor mij. Dankjewel!

Lieve **Henk & Rineke**, alle gezamenlijke etentjes en goede gesprekken in de afgelopen jaren waren een weldaad voor mij en ik hoop dat wij deze traditie nog lang voortzetten. Van harte dank voor alles!

Liebe **Mama** & Lieber **Burkhard**, Lieber **Papa** & Liebe **Marlies**, bei Euch liegt die Basis für alle Kraft, Selbstvertrauen und Unabhängigkeit, die mein Leben so viel schöner und promovieren ein Stück leichter machen. Danke für all das was ihr mir (mit auf den Weg) gebt!

Liebe Schwestern, **Debora, Janna** und **Rosa**, was wäre das Leben leer und langweilig ohne Euch. Danke für Euer Vertrauen, Eure Ehrlichkeit und Eure selbstverständliche und unendliche Unterstützung in allen Lebenslagen. Zu wissen dass ihr, Debora und Rosa, am 7. September neben mir steht gibt mir die nötige Zuversicht. Danke dafür!

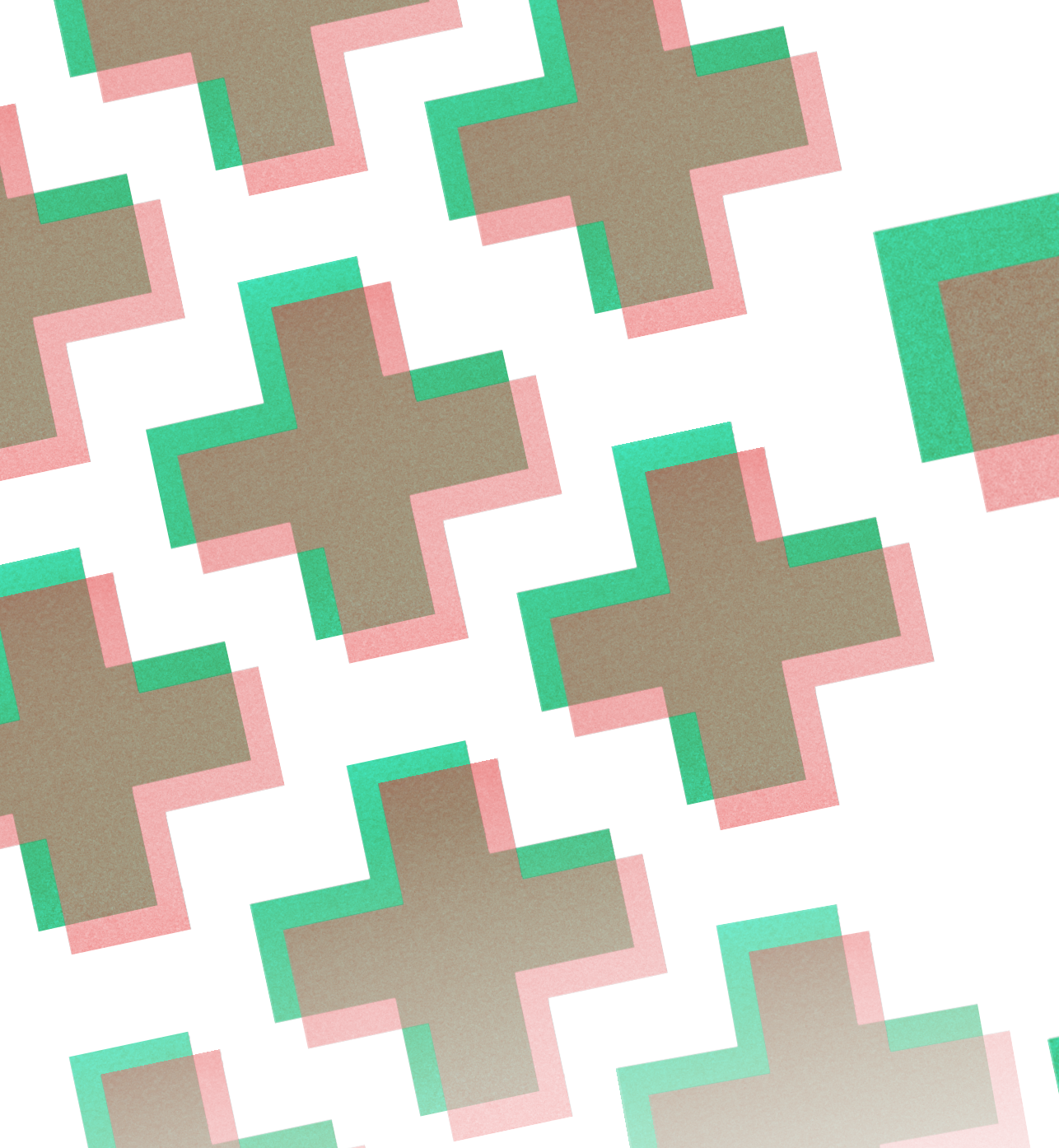
Sicco, jouw grenzeloze en onvoorwaardelijk vertrouwen en steun zijn voor mij onontbeerlijk en waren een drijvende kracht voor het afronden van dit onderzoek. Jij maakt mijn leven zo veel leuker! Dankjewel!

C

Curriculum vitae

CURRICULUM VITAE

Fania Ruth Gärtner was born on the 20th of August 1981, in Berlin, Germany. In 2001 she graduated from the 'Luise-Henriette Oberschule (Gymnasium)' in Berlin. Subsequently, she moved to the Netherlands to study Psychology at the University of Groningen. As part of her study, Fania spent four months at the Aston Business School, Aston University, Birmingham, United Kingdom. She obtained her Master's degree in Social Psychology in July 2006. After her graduation she was employed at the University of Groningen, where she coordinated and taught communication and consultation skills courses for psychology and medical students. From 2008 to 2012 she worked as a PhD student at the Coronel Institute for Occupational Health, Academic Medical Center (AMC), University of Amsterdam. She conducted research on the development of a measurement instrument and prevention strategies for impairments in work functioning due to common mental disorders with a focus on health care workers. During her time at the AMC, Fania was an active member of the works council and of the Network for Qualitative Studies in Health Research. Since April 2012 she works as a post-doctoral researcher at the Department of Medical Decision Making of the Leiden University Medical Center, the Netherlands. The focus of her work is on the development and evaluation of patient reported outcomes.



Workers in the health care sector are at high risk for developing mental health complaints which can have negative consequences for the quality of care provided. To help health care workers remain healthy and functioning well, preventive actions are essential.

This thesis describes several studies on the development and clinimetric evaluation of a questionnaire to assess work functioning impairments in nurses and allied health professionals: the ‘Nurses Work Functioning Questionnaire’.

Furthermore, the effectiveness of a newly developed workers’ health surveillance mental module was studied in a cluster randomized controlled trial. This preventive strategy combines online screening, including the new questionnaire, with an invitation for a preventive consultation with an occupational physician. The effects on help-seeking behavior, mental health complaints, and work functioning were examined.