Mental health symptoms as prognostic risk markers of all-cause and psychiatric sickness absence in office workers

Corné A. M. Roelen1,2, Rob Hoedeman1, Willem van Rhenen1,3, Johan W. Groothoff2, Jac J. L. van der Klink2, Ute Bültmann2

1 365Health Service, Utrecht, The Netherlands
2 Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
3 Center for Human Resource, Organization and Management Effectiveness, Business University Nyenrode, Breukelen, The Netherlands

Correspondence: Corné A. M. Roelen, 365ArboNed Corporate Accounts, PO Box 158, 8000 AD Zwolle, The Netherlands, tel: +31 38 4554700, fax: +31 38 4537272, e-mail: corne.roelen@arboned.nl

Background: To investigate mental health symptoms as prognostic risk markers of all-cause and psychiatric sickness absence (SA). Methods: Mental health symptoms were measured in 1137 office workers with the Four-Dimensional Symptom Questionnaire (4DSQ), including scales for distress, depression, anxiety and somatization. The total number of SA days was accumulated prospectively on the individual level and high SA was defined as ≥30 SA days during 1-year follow-up. Psychiatric SA was also tallied on the individual level during 1-year follow-up. Baseline 4DSQ scores were associated with high all-cause SA and psychiatric SA by logistic regression analysis. The Hosmer–Lemeshow test and calibration slope were used to assess the accuracy of predictions by 4DSQ scores. The ability of 4DSQ scores to discriminate high-risk from low-risk employees was estimated by the area under the receiver operating characteristic curve. Results: Six hundred thirty-three office workers (56%) participated in the study. All 4DSQ scales were prospectively associated with high all-cause SA and psychiatric SA. Distress and somatization scores showed acceptable calibration, but failed to discriminate between office workers with and without high all-cause SA. The distress scale did show adequate calibration (calibration slope = 0.95) and discrimination (area under the receiver operating characteristic curve = 0.71) for psychiatric SA. Conclusion: Distress was a valid prognostic risk marker for identifying office workers at work, but at risk of future psychiatric SA. Further research is necessary to investigate the prognostic performance of distress as risk marker of psychiatric SA in other working populations and to determine cut-off points for distress.

Introduction

Mental health problems are common in the European workforce1-3 and an increasing cause for sickness absence (SA)4-6 and work disability.7-13 Psychological distress is one of the most studied mental health problems and was recognized as a risk factor for the onset of long-term SA in Norway,12 the Netherlands,14 Finland15 and Sweden.16 Other mental health problems, such as depression and anxiety, are also highly prevalent in the workforce17-19 and related to long-term SA and disability pensioning.20,21 In the Dutch working population, a physician’s diagnosis of depression or anxiety disorder was associated with longer SA duration.22 Given the duration of SA due to mental health problems, the risk of disability pensioning, the costs for employers and society and the marginalization of long-term sick-listed employees from the labour market, it is important to identify high-risk employees at an early stage, i.e. before long-term SA occurs.23 Recently, Stansfeld et al.23 showed that clinical, but not sub-threshold, common mental disorders were associated with an increased risk of psychiatric SA in male, but not female, civil servants of the British Whitehall II Study. The authors recommended that public and occupational health services should focus on the identification of employees with mental health problems, as they might be at risk of future psychiatric SA.

The aforementioned studies have only reported associations between mental health symptoms and long-term (psychiatric) SA. Associations neither tell us how accurately mental health symptoms predict future SA (calibration) nor if mental health symptoms distinguish between employees with and without long-term SA (discrimination). This study calibrated mental health symptoms as prognostic risk markers of all-cause and psychiatric SA and investigated their discriminative ability.

Methods

At baseline, all 1137 employees working at an insurance office were invited to participate in an occupational health check-up, consisting of a questionnaire measuring general health, mental health, work environment, commitment to work and coping styles. Employees were referred to an occupational physician (OP) if the questionnaire results necessitated preventive consultation or treatment. Baseline mental health scores measured in the health check-ups were linked to individual SA in the following year. Ethical approval for this study was granted by the Medical Ethics Committee of the University Medical Center Groningen.

Mental health symptoms

The Four-Dimensional Symptom Questionnaire (4DSQ) is a valid and reliable 50-item self-report questionnaire for measuring distress, depression, anxiety and somatization in primary and occupational healthcare settings.24,25 The distress scale consists of 16 items (score range 0–32) addressing symptoms elicited by a stressor or the efforts to maintain psychosocial functioning (Cronbach α = 0.90), e.g. worry, irritability, tension, listlessness, poor
concentration, sleeping problems and demoralization. The depression scale contains six items (score range 0–12) about loss of pleasure and depressive thoughts including suicidal ideation ($\alpha = 0.82$). The anxiety scale consists of 12 items (score range 0–24) about panic attacks, phobic anxiety, free-floating anxiety and avoidance behaviour ($\alpha = 0.77$). The somatization scale comprises 16 medically unexplained physical symptoms (score range 0–32), such as dizziness, upset stomach, palpitations and muscular aches ($\alpha = 0.78$). Employees could indicate how often they experienced these symptoms in the last week in the categories ‘no’, ‘sometimes’, ‘regularly’, ‘often’ and ‘very often or constantly’. For scale scores, the responses were categorized as 0 for ‘no’, 1 for ‘sometimes’ and 2 for the other response categories. The scale scores were standardized as percentage of the maximum scale score, so that the strengths of associations could be compared across 4DSQ scales.

**Sickness absence**

SA was defined as absence from work due to any (i.e. work-related and non-work-related) injury or illness. The office workers reported sick to their manager who sent the sick report to the occupational health service (OHS) on the first SA day. The day an employee resumed work, the manager sent a recovery report to the OHS. The calendar days between the first and last day of SA were counted as SA days, with partial days off work treated as full SA days. The number of SA days was accumulated for each employee during 1-year follow-up and was considered high when in the upper decile of the SA distribution. The 90th percentile was 30 SA days and therefore high SA was defined as 30 SA days. It should be acknowledged that these were accumulated and not necessarily consecutive SA days.

In The Netherlands, SA is medically certified by an OP in the third or fourth SA week. The OP records the SA diagnosis in the OHS register according to the 10th version of the International Classification of Diseases (ICD-10). Psychiatric SA was defined as SA certified by the OP with an ICD-10 diagnostic code F00–F99, i.e. Chapter V ‘Mental and behavioural disorders’ of ICD-10. Psychiatric SA included specific psychiatric morbidities, such as schizophrenia, unipolar and bipolar depressions and personality disorders, as well as less well-specified morbidities, such as stress-related, neurotic and somatoform disorders.

**Prediction modelling and performance of risk markers**

The baseline 4DSQ scale scores were associated with high all-cause SA and psychiatric SA by logistic regression modelling. As we were interested in the performance of the 4DSQ scales as prognostic risk markers of SA and not in causal relationships, we did not adjust the logistic regression analyses for potential confounders. The results of logistic regression are shown as odds ratios with 95% confidence intervals.

The prognostic performance of risk markers was addressed in terms of calibration and discrimination. Calibration refers to the agreement between the probabilities of high all-cause SA or psychiatric SA predicted by 4DSQ scores and the observed probabilities. Calibration was performed with the Hosmer–Lemeshow goodness-of-fit test and was considered adequate for Hosmer–Lemeshow $P \geq 0.05$. Discrimination refers to the ability of 4DSQ scales to distinguish employees with high all-cause SA or psychiatric SA from those without high all-cause SA or psychiatric SA. The area under the receiver operating characteristic curve (AUC) was regarded as a measure for discrimination. An AUC of 0.5 indicates no discrimination above chance and an AUC of 1.0 indicates perfect discrimination. A rough guide for classifying the discriminative ability is AUC <0.6 fail, AUC 0.6–0.7 poor, AUC 0.7–0.8 fair, AUC 0.8–0.9 good and AUC 0.9–1.0 excellent. In this study, discrimination was considered adequate if AUC $\geq 0.70$.

**Internal validation of risk markers**

A prognostic risk marker will perform better in the sample used to assess the risk marker than in new populations. To correct for this over-optimism, the 4DSQ scales with adequate calibration and discrimination were validated by bootstrapping in R by using Harrell’s Regression Modeling Strategies package, version 3.2-0. Bootstrapping is a data simulation technique in which subjects are drawn at random from the study population with replacement, i.e. one subject can be drawn into the bootstrap sample several times. Bootstrap samples were of equal size as the original sample, but had different data structures. The difference in performance between the bootstrap samples and the original sample was used to correct the performance parameters. Over-optimism-corrected calibration was expressed as the slope of the calibration plot, a scatter graph plotting predicted probabilities against observed probabilities. The discriminative ability was validated by shrinking the AUC by the amount of over-optimism.

**Results**

Of the 1137 office workers, 633 (56%) participated in the health check-ups and agreed that their results were linked to SA data accumulated during 1-year follow-up. Table 1 shows the characteristics of the study population. Non-participants were younger and more likely to be male than participants. Among participants, employees working in client service jobs were underrepresented and those working in information and communication technology were overrepresented. The number of SA days during 1-year follow-up did not differ significantly between participants and non-participants.

**Mental health scores**

The non-standardized 4DSQ scores of participants averaged 4.47 (SD = 5.44) for distress, 0.31 (SD = 1.12) for depression, 0.54 (SD = 1.30) for anxiety and 4.16 (SD = 4.34) for somatization. There were no significant age and gender differences (table 2). Employees in information and communication technology reported higher anxiety scores than employees in other jobs (chi-square $P < 0.01$). Employees working in client service reported higher somatization scores (chi-square $P < 0.01$).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants, N (%)</th>
<th>Non-participants, N (%)</th>
<th>Chi-square analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;35</td>
<td>99 (17)</td>
<td>97 (22)</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td>35–44</td>
<td>203 (34)</td>
<td>179 (40)</td>
<td></td>
</tr>
<tr>
<td>45–54</td>
<td>189 (32)</td>
<td>126 (28)</td>
<td></td>
</tr>
<tr>
<td>$\geq 55$</td>
<td>107 (18)</td>
<td>47 (11)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>390 (62)</td>
<td>341 (68)</td>
<td>$P = 0.04$</td>
</tr>
<tr>
<td>Women</td>
<td>243 (38)</td>
<td>163 (32)</td>
<td></td>
</tr>
<tr>
<td>Job</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client service</td>
<td>158 (27)</td>
<td>158 (38)</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td>Administration and facilities</td>
<td>177 (31)</td>
<td>148 (36)</td>
<td></td>
</tr>
<tr>
<td>ICT and applications</td>
<td>152 (26)</td>
<td>25 (6)</td>
<td></td>
</tr>
<tr>
<td>Managers</td>
<td>55 (9)</td>
<td>42 (10)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>37 (6)</td>
<td>38 (9)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>54</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>SA days during follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>257 (41)</td>
<td>182 (36)</td>
<td>$P = 0.30$</td>
</tr>
<tr>
<td>1–9</td>
<td>218 (34)</td>
<td>170 (34)</td>
<td></td>
</tr>
<tr>
<td>10–19</td>
<td>75 (12)</td>
<td>75 (15)</td>
<td></td>
</tr>
<tr>
<td>20–29</td>
<td>24 (4)</td>
<td>26 (5)</td>
<td></td>
</tr>
<tr>
<td>$\geq 30$</td>
<td>59 (9)</td>
<td>51 (10)</td>
<td></td>
</tr>
</tbody>
</table>

ICT, information and communication technology.
A total of 513 (81%) employees reported good general health and 27 (4%) reported a health condition. Employees with existing health conditions had higher scores on all mental health symptoms (table 2).

During 1-year follow-up, 59 office workers (9%) had high all-cause SA. The standardized distress, depression, anxiety and somatization scores at baseline were positively associated with future high all-cause SA during follow-up, with somatization and depression as strongest predictors (table 3). In 39 (66%) office workers, high SA was due to psychiatric causes. The odds of psychiatric SA in the following year. Calibration was adequate for the performance of the 4DSQ scales as risk markers of SA.

The performance of this scale was internally validated in 200 bootstrap samples. Distress scores were missing in 15 cases (2%); the remaining 618 office workers with complete data were used for bootstrapping. The bootstrapped calibration slope = 0.95, indicating a 5% over-optimism of predictions. The over-optimism-adjusted discriminative ability was AUC = 0.71, indicating that distress will correctly discriminate white-collar workers at high risk of psychiatric SA from those at low risk in 71% of cases.

### Discussion

Distress, depression, anxiety and somatization scales at baseline were prospectively associated with both high all-cause SA and psychiatric SA in the following year. Calibration was adequate for the distress and somatization scales, but these scales did not adequately distinguish between office workers with and without high all-cause SA. Distress did discriminate office workers with psychiatric SA from those without psychiatric SA and proved to be a valid prognostic risk marker of future psychiatric SA in white-collar populations.

### Mental symptoms as risk markers of all-cause SA

The study confirmed the previously described associations of mental health symptoms with long-term SA, although contrasted the results of psychiatric SA from those without psychiatric SA. Therefore, the study confirmed the previously described associations of mental health symptoms with long-term SA, although contrasted the results.
of Stansfeld et al., who found that common mental disorders were not related to all-cause SA. The association between mental health symptoms and all-cause SA in our study might be explained by shared variance, as 66% of high all-cause SA was due to psychiatric disorders. Stansfeld et al. assessed the frequency of long-term (>7 days) SA episodes, while we investigated accumulated SA days. Furthermore, Stansfeld et al. addressed mental health problems with the General Health Questionnaire, while we measured different types of mental health symptoms with the 4DSQ, an instrument generally used in Dutch primary and occupational health care.4,25

Somatization and depression were the strongest predictors of high all-cause SA in agreement with the bivariate results of Terluin et al.60 Although associated with future all-cause SA, none of the 4DSQ scales was a valid prognostic risk marker for high all-cause SA. The 4DSQ depression and anxiety scales showed miscalibration, which meant that they did not accurately predict the risk of high all-cause SA. This may be due to the low prevalence of depressive and anxiety symptoms in the study population. An earlier study of employees working in a Dutch telecom company also showed low prevalences of depression.24 Possibly, the 4DSQ depression scale measures severe depression, while more frequent mild depressions are reflected in higher distress scores.25 The 4DSQ distress and somatization scales showed adequate calibration, but failed to discriminate office workers with high all-cause SA from those without high all-cause SA.

Mental symptoms as risk markers of psychiatric SA

All 4DSQ scales were associated with a higher risk of future psychiatric SA, but the 4DSQ depression and anxiety scales showed miscalibration. The 4DSQ somatization scale showed adequate calibration, but poorly discriminated between office workers with and without psychiatric SA. The 4DSQ distress scale showed adequate calibration and discrimination as risk marker of psychiatric SA. It should be acknowledged here that the majority of psychiatric SA in The Netherlands is caused by stress-related and neurotic disorders.31 After internal validation and correction for over-optimism, the 4DSQ distress scale still showed fair discrimination (AUC = 0.71), although an AUC ≥0.75 is considered as practically useful.12

Strengths and limitations

An asset of the study is the use of recorded SA data instead of potentially recall-biased self-reported SA.3,34 Furthermore, misclassification of 4DSQ scores and SA was minimized by the prospective design and the different data sources (questionnaire for mental health symptoms and OHS register for SA) limited common method bias. Although the study population was a convenience sample of office workers, the over-optimism-adjusted calibration and discrimination reflect the prognostic performance that is to be expected in other white-collar populations. A limitation of the study is that the 4DSQ is virtually only used in The Netherlands, which restricts international comparisons.

The health check-ups may have increased the employees’ awareness of health risks and may have affected their health behaviour, which might, subsequently, have reduced SA. This could have underestimated associations between mental health symptoms and SA, but did not impact calibration and discrimination of distress as prognostic risk marker. We included medically certified psychiatric SA, which is preferred to employee-reported mental health problems, but may have underestimated the number of psychiatric SA episodes as SA was OP-certified in the third or fourth SA week.

The participating employees were informed that the health check-up results would be linked to their SA in the following year, but they did not know which health check-up data were used, thus preventing under-participation of employees with mental health symptoms. Still, 56% of all office workers working at the insurance office participated in the health check-ups and the effective sample sizes were N = 59 for high all-cause SA and N = 39 for psychiatric SA. Each logistic regression model included only one variable (i.e. one of the 4DSQ scales) at a time. Regression models are assumed to provide stable predictions when there are 10 or more events per variable.27,35 This assumption was amply met in this study.

Twenty-seven participants reported an existing health condition, but information on diagnoses was not available from the health check-ups. Existing physical disorders might lead to over-optimistic predictions of high all-cause SA. However, mental health symptoms failed as risk markers of all-cause SA. Individuals with existing mental disorders are likely to report more mental health symptoms and have an increased risk of psychiatric SA. Although necessary for examining independent effects of mental health symptoms on SA, adjustment for existing health conditions is not required when investigating mental health symptoms as prognostic risk markers of SA.

Practical implications and future directions

Given the public health impact of mental health symptoms on occupational outcomes,3 public and occupational health services should focus on the early identification of employees with mental health problems.33 This study investigated mental health symptoms as prognostic risk markers of SA and found that distress was a valid risk marker of psychiatric SA in office workers. Hence, measuring distress levels could early identify office workers at risk of psychiatric SA. However, it is too premature to recommend the use of distress to identify employees at risk of psychiatric SA in occupational and public healthcare practice. Bootstrapping techniques validated distress as predictor of future psychiatric SA for white-collar populations, but its prognostic performance in other (e.g. blue-collar) working populations and the outcome of screening working populations for distress still need to be investigated. A randomized controlled trial by Wang et al.36 showed that screening for depression, followed by a systematic programme of telephone outreach and care management, decreased symptoms and increased job retention and hours worked. Alternatively, there is evidence from medical settings that screening for depression is not associated with improved outcomes.37

Furthermore, care must be taken as screening for mental health symptoms may raise awareness and herewith medicalize transient distress symptoms. Therefore, it is important to establish cut-off points for distress that specifically identify high-risk employees, otherwise too many employees would be unnecessarily invited for medical consultations. To determine these specific cut-off points, research in larger and more heterogeneous working populations is required.

If distress or other mental health symptoms accurately identify employees at work, but at risk of psychiatric SA, then occupational and public healthcare providers might consider screening working populations for mental health symptoms. In most European countries, health and work conditions are assessed on a regular basis by questionnaire surveys or health check-ups. An advantage of such surveys or check-ups is that employees are used to them, so they do not raise awareness and do not pose a risk of stigmatization or medicalization. However, a problem is the ‘healthy volunteer effect’, i.e. healthy employees are more likely to participate than employees with health problems.38,39 Thus, surveys or health check-ups may not get the high-risk employees.

Occupational healthcare providers and managers share a responsibility to motivate as many employees as possible to participate in health surveys, especially employees with health complaints. When we know which mental health symptoms identify employees at risk of psychiatric SA, managers and supervisors could be trained in recognizing these symptoms. Furthermore, digital applications could be developed to collect employee-reported health symptoms.
Dale and Hagen (2007) concluded that digital personal assistants performed better than ‘pen and paper’ methods, although technical malfunctions need to be dealt with and more research is necessary to compare digitally collected data with paper data. Applications could present employees with their risk of long-term (psychiatric) SA and advise high-risk employees to consult occupational health providers who can assess the need for further diagnosis and early treatment.

Conflicts of interest: None declared.

Key points
- Mental health problems are prospectively associated with SA, but so far mental health symptoms have not been validated for their use as prognostic risk markers.
- This study showed that mental health symptoms failed as prognostic risk markers of all-cause SA in office workers.
- Psychological distress was a valid prognostic risk marker of psychiatric SA in office workers.

References
29 Harrell FE. Regression modelling strategies. Available at: http://cran.r-project.org/web/packages/rms/ (13 April 2012, date last accessed).