

# BMJ Open Cross-national validation of the MHQoL: psychometric evaluation and open-source tools for assessing mental health quality of life

Stijn B Peeters <sup>1</sup>, Frederick W Thielen,<sup>2</sup> Marleen De Mul <sup>3</sup>,  
Marjo Sinokki <sup>4,5</sup>, Beatriz Olaya,<sup>6,7,8</sup>  
Christina Maria Van Der Feltz-Cornelis <sup>9,10</sup>, Leona Hakkaart-Van Roijen<sup>3</sup>

**To cite:** Peeters SB, Thielen FW, De Mul M, *et al*. Cross-national validation of the MHQoL: psychometric evaluation and open-source tools for assessing mental health quality of life. *BMJ Open* 2026;**16**:e108598. doi:10.1136/bmjopen-2025-108598

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2025-108598>).

Received 29 July 2025  
Accepted 21 April 2026



© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY. Published by BMJ Group.

For numbered affiliations see end of article.

## Correspondence to

Stijn B Peeters;  
s.b.peeters@eshpm.eur.nl

## ABSTRACT

**Objectives** To validate the cross-national psychometric properties of the Mental Health Quality of Life questionnaire (MHQoL) and to develop an open-source toolbox for its scoring, transformation and presentation.

**Design** Secondary analysis of data from a multicentre international randomised controlled trial (EMPOWER).

**Setting** Workplace settings in small-sized and medium-sized enterprises (SMEs) and public sector organisations in Finland, Spain and the UK.

**Participants** The sample included 564 employees: 122 from Finland, 114 from Spain and 328 from the UK. Most were white-collar workers in SMEs or public organisations, mainly in public administration, manufacturing, health/life sciences or higher education. Women were the majority (56%–91% across countries), and mean age ranged from 43 to 48 years.

**Interventions** No intervention was delivered for this analysis; data were drawn from baseline assessments.

**Primary and secondary outcome measures** Primary outcomes were internal consistency and construct validity of the MHQoL, evaluated using Cronbach's alpha, measurement invariance testing and multilevel analyses of associations between MHQoL dimensions and its visual analogue scale (VAS). Secondary outcomes were convergent validity, assessed through correlations between MHQoL scores and other mental health and quality of life measures (EuroQol 5-Dimension 5-level questionnaire (EQ-5D-5L), Patient Health Questionnaire-9 (PHQ-9), Generalized Anxiety Disorder-7 (GAD-7), Insomnia Severity Index, Perceived Stress Scale-4 (PSS-4), Psychosocial Risk Scale, and World Health Organization Five Well-Being Index (WHO-5)).

**Results** The MHQoL showed good internal consistency across countries, with Cronbach's alpha ranging from 0.741 in Finland to 0.806 in Spain (overall  $\alpha=0.787$ ). Measurement invariance across Finland, Spain and the UK supported construct validity. Multilevel regression analyses showed associations between MHQoL dimensions and the MHQoL-VAS, with strongest contributions from Self-Image, Daily Activities, Mood and Future. Convergent validity was supported by moderate to strong correlations between MHQoL, EQ-5D-5L and related mental health measures. An open-source R package and Shiny web application ('MHQoL Toolbox') were developed for scoring, transformation and visualisation

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Inclusion of a multicountry dataset enabled cross-national comparisons and accounted for cultural differences in workplace environments.
- ⇒ Reliability and validity of the Mental Health Quality of Life (MHQoL) were assessed using multiple psychometric approaches, including internal consistency, construct validity and convergent validity.
- ⇒ The study combined psychometric validation with the development of an open-source toolbox to support standardised scoring, transformation and presentation of MHQoL data.
- ⇒ Participants were predominantly female, highly educated and employed in public sector or white-collar roles, which may limit the generalisability of findings to other worker populations.
- ⇒ The sample was unevenly distributed across countries and no country-specific EQ-5D-5L value set was available for Finland, requiring the use of a proxy.

**Conclusions** The MHQoL is a reliable and valid measure of mental health-related quality of life across countries. The MHQoL toolbox supports consistent, transparent implementation, facilitating use in research, clinical practice and economic evaluations.

**Trial registration number** NCT04907604.

## INTRODUCTION

Mental health disorders are increasingly prevalent in modern society, with global figures continuing to rise. In 2019, an estimated 970 million people, equivalent to one in every eight individuals, were living with a mental disorder worldwide.<sup>1</sup> This number has risen significantly in recent years. Preliminary estimates suggest a 26% increase in anxiety disorders and a 28% rise in major depressive disorders within just 1 year.<sup>2</sup> The societal economic burden associated with mental illness on society is substantial, with



annual costs ranging from €54 million to €280 billion per country,<sup>3</sup> totalling approximately €450 billion across the European Union (EU)<sup>4</sup> and US\$2.5 trillion worldwide.<sup>5</sup>

The workplace is also significantly affected by these disorders. Each year, 27% of EU workers report work-related stress, depression or anxiety.<sup>6</sup> These conditions contribute to absenteeism, which is more affected by mental health issues than physical conditions,<sup>6-9</sup> reduced job performance (presenteeism),<sup>7 10</sup> lower Quality of Life (QoL),<sup>10-13</sup> higher medical costs<sup>10</sup> and employment difficulties.<sup>11</sup> The resulting economic burden is substantial, with depression and anxiety leading to 12 billion lost workdays globally each year and an estimated €900 billion loss in productivity.<sup>11</sup>

To improve QoL and reduce the economic burden on society, interventions are required to enhance QoL at a reasonable expense. Economic evaluations can help identify treatments that provide the best value for money. Hence, it is essential that both costs and QoL are measured accurately to ensure a valid representation. Commonly used questionnaires like the 36-item Short-Form Health Survey (SF-36)<sup>14</sup> and the five-dimension EuroQol questionnaire (EQ-5D)<sup>15</sup> are, however, primarily focused on physical aspects and may not sufficiently capture the domains of mental health problems in their assessments of QoL. As a result, the impact of mental health on QoL is less represented in the outcomes of these questionnaires.<sup>16-18</sup>

To overcome this, the Mental Health Quality of Life (MHQoL)<sup>19</sup> questionnaire was developed, which is a preference-based instrument designed to assess QoL using dimensions that individuals with mental health challenges identify as important and value highly. These dimensions include Self-Image, Independence, Mood, Relationships, Daily Activities, Physical Health and Hope. Hence, the MHQoL can be used alongside generic measures in economic evaluation studies such as cost-effectiveness analysis.

While previous psychometric evaluations of the MHQoL have been conducted,<sup>19 20</sup> they did not specifically target the working population. However, this focus is crucial given the unique characteristics of the demographic, including the dual pressures of professional and personal responsibilities, exposure to work-related stressors and the impact of mental health on productivity and absenteeism. Mental health problems among working-age individuals are a growing concern in both public health and economic terms, contributing significantly to societal costs through reduced labour participation and increased healthcare use. Valid and practical tools to measure mental health-related QoL in this group are therefore essential to inform targeted interventions, workplace policies and economic evaluations.<sup>21</sup> Furthermore, no studies have compared the MHQoL and the EQ-5D-5L in conjunction with other mental health questionnaires (eg, anxiety, depression, stress, work stress and well-being questionnaires), or cultural differences between

countries. Understanding these differences is important because it provides valuable insights into how each questionnaire addresses mental health issues. Moreover, there is not a standardised tool for calculating the MHQoL. Although a manual is available, calculations must be performed manually each time, making the process time-consuming and prone to inconsistencies and errors.

The objectives of this study were twofold. First, to assess the psychometric performance of the MHQoL including reliability and validity, when applied in a working population across Finland, Spain and the United Kingdom, and to compare these findings with prior evidence from the general population by Enzing *et al.*<sup>20</sup> Second, to develop a standardised, open-access toolbox for scoring, transforming and reporting MHQoL outcomes. This toolbox aimed to promote reproducibility, user-friendliness and wider uptake of the MHQoL in research and health economic practice, among users and health policy makers.

## METHODS

### Data sources

Data for this study were derived from the EMPOWER trial (trial number NCT04907604), which has been described in detail elsewhere.<sup>22</sup> The EMPOWER trial was a multi-centre cluster randomised controlled trial conducted in workplace settings, in which organisations (clusters) were recruited and employees were invited to participate on a voluntary basis. Eligible participants were employees working in participating organisations who had access to the eHealth platform and were able to complete questionnaires in the respective country language. Recruitment was conducted through participating organisations, where employees received information about the study and were invited to enrol.

For the present analysis, cross-sectional baseline data from the trial was used. Preparation of this report was guided by the Consolidated Standards of Reporting Trials (CONSORT) checklist<sup>23</sup>; however, as this is a secondary analysis of baseline data from a multicentre international randomised controlled trial, full details of the trial design and methodology are reported elsewhere,<sup>22</sup> and several CONSORT items were therefore not applicable.

### Population

The sample included 564 employees, derived from baseline data of the EMPOWER trial and based on the availability of complete baseline questionnaire data. While the original EMPOWER study aimed to recruit 729 participants to evaluate intervention effectiveness,<sup>22</sup> the present study focuses on the psychometric evaluation of the MHQoL; therefore, this sample size was considered adequate for the present analyses. Of these participants, 122 were from Finland, 114 from Spain and 328 from the UK. Female participants predominated (56.1% in Spain, 91.0% in Finland, 82.9% in the UK). The mean age was 46.5 years in Finland (range: 21–67), 47.8 years in Spain

(range 21–66) and 43.1 years in the UK (range: 23–72). Most participants were white-collar workers (~90%) and employed in the public sector, reaching 100% in Finland and the UK. In Finland, all participants worked in public administration, in Spain, the majority were in manufacturing (59.6%), with others in public administration; in the UK, most were employed in health/life sciences (82.9%) and higher education. See online supplemental appendix A, table 1 for full details.

### Instruments and outcome measures

Participants completed self-report questionnaires via the EMPOWER eHealth platform, using validated Finnish, Spanish and English versions of the instruments. The EQ-5D-5L language versions were obtained from the EuroQol Group, while the MHQoL translations were provided by Erasmus University Rotterdam.

The MHQoL includes seven dimensions (i.e. self-image, independence, mood, relationships, daily activities, physical health and hope) on a four level scale (0=very poor, 3=very good least favourable condition and 3 the most favourable.<sup>19</sup> The total score ranges from 0 to 21, with higher scores indicating better mental health-related QoL. It also includes a visual analogue scale (MHQoL-VAS) from 0 (worst imaginable psychological well-being) to 10 (best). Preference-based utilities are available only for the Dutch dataset and therefore not used in the statistical analyses.

The EQ-5D-5L is a standardised instrument for assessing health-related QoL, covering five dimensions: mobility, self-care, daily activities, pain/discomfort and anxiety/depression.<sup>24</sup> Each is rated on five levels of severity. Results include preference-based utility scores, Level Sum Scores (LSS, range: 5=best health to 25=worst health),<sup>25</sup> and a VAS (0=worst, 100=best health).<sup>24</sup> Country-specific value sets were applied<sup>26 27</sup>; as no Finnish value set was available, Denmark's<sup>28</sup> was used as best proxy.

The PHQ-9 is a self-assessment protocol to assess severity of depression via nine items scored 0–3, yielding a total score from 0 to 27<sup>29</sup>. Scores indicate severity levels: minimal (1–4), mild (5–9), moderate (10–14), moderately severe (15–19) and severe (20–27).

The GAD-7 is a 7-item scale for identifying Generalised Anxiety Disorder and assessing symptom severity.<sup>30</sup> Each item is rated 0–3, with total scores from 0 (no symptoms) to 21 (severe symptoms).

The Insomnia Severity Index (ISI) is a 7-item tool evaluating issues related to falling and staying asleep, satisfaction with sleep, daytime impairment and distress about sleep problems.<sup>31</sup> Each item is scored 0–4, with a total score of 0–28. Higher scores indicate more severe insomnia.<sup>32</sup>

The PSS-4 is a 4-item instrument measuring perceived stress.<sup>33</sup> Items are rated 0 ('never') to 4 ('very often'), yielding a score from 0 to 16, with higher scores indicating greater stress.

The Psychosocial Risk Scale (PRS) measures work-related stress.<sup>34</sup> It includes 16 items scored from 0 ('not at

all') to 3 ('very stressed'), resulting in a score from 0 (no stress) to 48 (extreme stress).

The WHO-5 is a brief self-assessment for current mental wellbeing.<sup>35</sup> Respondents rate five items from 0 ('at no time') to 5 ('all of the time'). The raw score (0–25) is multiplied by 5 to give a final score (0–100), where indicates worst and 100 best possible well-being.

### Statistical analysis

The statistical analysis was structured in multiple parts. First, descriptive statistics provided an overview of the mean score per person based on completed (mental) health questionnaires. A detailed analysis of MHQoL dimensions was conducted, reporting mean scores per dimension along with standard errors, medians and minimum-maximum values.

Second, Cronbach's alpha and McDonald's omega ( $\omega$ ), along with their 95% CIs were calculated for the seven MHQoL dimensions to assess internal consistency overall and per country. Internal consistency reflects the extent to which test items measure the same construct.<sup>36</sup> Cronbach's alpha assumes tau-equivalence, whereas McDonald's omega provides a more flexible estimate of reliability that does not rely on this assumption. Alpha values range from 0 (no correlation) to 1 (full correlation), with values over 0.70<sup>36</sup>.

Third, to examine construct validity, measuring invariance of the MHQoL across countries was tested using multi-group confirmatory factor analysis (CFA). Invariance was assessed hierarchically<sup>37</sup> starting with configural invariance (same pattern of factor loadings across groups), followed by metric invariance (statistically identical factor loadings) and scalar invariance (statistically equal thresholds).<sup>38</sup> Given the ordinal scale, Weighted Least Squares with Mean and Variance adjustments were applied,<sup>38</sup> using thresholds instead of intercepts while residual invariance was not assessed.<sup>38</sup> Model fit was evaluated using the Comparative Fit Index (CFI), root mean square error of approximation (RMSEA) and standardised root mean square residual (SRMR), with invariance thresholds set at CFI $\geq$ 0.95, RMSEA $\leq$ 0.06 and SRMR $\leq$ 0.08.<sup>39</sup> A  $\chi^2$  difference test was conducted to assess significant differences in model fit between levels of invariance. Besides, Spearman's rank correlation coefficients were calculated to assess associations between: MHQoL and MHQoL VAS; EQ-5D-5L preference-based utilities and LSS with MHQoL and MHQoL VAS at both overall and country levels. Spearman's rank correlation was chosen over Pearson's R due to the ordinal nature of MHQoL dimensions.<sup>40</sup> The correlation (Spearman's  $r$ ) between EQ-5D-5L and MHQoL scores was expected to be negative, as better (mental) QoL results in lower EQ-5D-5L sum scores but higher MHQoL scores. Conversely, EQ-5D-5L utilities were expected to show positive correlations, as higher utility scores indicate better QoL. Lastly, two multilevel regression analyses were conducted to account for the hierarchical structure of the data, with individuals nested within countries. The



first analysis examined the association between MHQoL dimensions and MHQoL-VAS, while the second assessed the impact of EQ-5D-5L dimensions on MHQoL to identify health-related QoL factors influencing mental QoL. Both models included fixed effects for predictors, along with individual-level covariates (age and gender (in this study the term gender is used to refer to sex assigned at birth)). Two model specifications were compared: (1) a random intercept model, allowing intercepts to vary across countries and (2) a random intercept and random slope model, allowing both intercepts and slopes to vary. A likelihood ratio test determined the best-fitting model, favouring the random intercept model. Model fit was evaluated using adjusted  $R^2$ , ranging from 0 (no fit) to 1 (perfect fit).<sup>41</sup>

To evaluate convergent validity, Spearman's rank correlation coefficients were calculated between the MHQoL and disease-specific mental health questionnaires (eg, PHQ-9, GAD-7, ISI, PSS-4, PRS, WHO-5), and likewise between the EQ-5D-5L and these questionnaires.

P-values were calculated, with significance set at  $p < 0.001$ . Spearman's rank correlation coefficients (Spearman's  $r$ ) were classified as very strong (0.90–1.00), strong (0.70–0.89), moderate (0.40–0.69), weak (0.10–0.39) and negligible (0.0–0.10).<sup>42</sup>

All the statistical analyses were performed using R V.4.3.2.<sup>43</sup>

### Development of the standardised MHQoL toolbox

The standardised MHQoL toolbox for transformation, calculation, and presentation of collected MHQoL responses was developed as both an R-package and R-Shiny application. The package allows for standardised data transformations and calculations by providing a set of functions in R. It also forms a foundational basis of the application and is based on the original MHQoL manual together with the Dutch value set by van Krugten *et al.*<sup>19 44</sup>

The R-Shiny application was built as an interactive tool to analyse, interpret, and visualise (up)loaded user data. The aim was to provide a graphical user interface for accessible result generation and presentation without the need for deeper knowledge of the statistical programming language R.

The application uses transformation functions from the R-package to convert, test, and score individual responses into scores and utility values, including reverse transformations. Calculation functions are implemented to compute both individual and average LSS and utility values for each respondent automatically.

Throughout development, one author documented all steps and implemented the function, after which a second author independently verified the outputs and usability of the toolbox by reproducing the calculations and checking the consistency of results across example datasets. This verification involved comparing the toolbox outputs with the results obtained to ensure that the implemented functions reproduced the expected results.

A list of R packages used both for the statistical analyses and in the development of the toolbox can be found in online supplemental appendix B.

### Patient and public involvement

Members of the public were involved in the development and design of the study, including the selection of relevant outcomes and the overall approach.<sup>22</sup> However, they were not involved in the conduct of the analyses or interpretation of the results presented in this paper.

## RESULTS

### Descriptives

In general, all participants scored low on stress, anxiety, depression and insomnia severity scales. Health-related QoL (EQ-5D-5L) showed favourable utility and LSS outcomes across all countries, with preference-based utilities aligning with expected values for a healthy population.<sup>45</sup> While MHQoL and MHQoL-VAS scores did not indicate poor mental QoL, they suggested more room for improvement than EQ-5D-5L. The UK consistently reported higher problem severity across all questionnaires. No country differences were found in MHQoL dimension-specific scores. All mean scores for (mental) health questionnaires and MHQoL dimensions are detailed in online supplemental tables 2 and 3 of appendix A.

### Internal consistency MHQoL

The MHQoL demonstrated strong internal consistency across all countries, with Cronbach's  $\alpha$  and McDonald's  $\omega$  generally exceeding 0.7. While the lower bound of the 95% CI for Cronbach's  $\alpha$  in Finland fell slightly below 0.7, McDonald's  $\omega$  remained above 0.7 across all countries. Additionally, a strong correlation was found between the MHQoL score and MHQoL-VAS in all countries (Spearman's  $r > 0.7$ ,  $p < 0.001$ ), confirming consistency between those two measures. Detailed results are presented in table 1.

### Construct validity

#### Measurement invariance test

The measurement invariance test (table 2) indicated good fit for the configural model, confirming a consistent factor loading pattern across countries. Metric invariance was supported, suggesting statistically equivalent factor loadings. However, the scalar model showed significant misfit compared with the metric model. To address this, thresholds for several items (Self-Image, Independence, Relationships and Physical Health) were freed, resulting in a partial scalar invariance model with acceptable fit indices. According to established guidelines, partial scalar invariance can still allow meaningful comparisons of latent factor means when a sufficient number of parameters remain invariant across groups.<sup>46</sup> Therefore, the partial scalar invariance model was retained for cross-country comparisons of the MHQoL. The specific

**Table 1** Internal consistency expressed in Cronbach's alpha of the seven MHQoL dimensions and Spearman's rank correlation coefficients expressed in Spearman's r for Finland, Spain, the UK and overall to present the correlation between the MHQoL score and MHQoL VAS

Country	Cronbach's alpha for the seven MHQoL dimensions*	McDonald's omega for the seven MHQoL dimensions*	Spearman's r to present the correlation between the MHQoL and MHQoL VAS
Finland (N=122)	0.741 (0.652 to 0.801)	0.768 (0.703 to 0.820)	0.702†
Spain (N=114)	0.806 (0.710 to 0.859)	0.814 (0.742 to 0.869)	0.767†
UK (N=328)	0.801 (0.761 to 0.835)	0.804 (0.768 to 0.836)	0.771†
Overall (N=564)	0.787 (0.755 to 0.815)	0.796 (0.765 to 0.822)	0.753†

\*Coefficient (95% CI).

†p<0.001

MHQoL, Mental Health Quality of Life; VAS, Visual Analogue Scale.

thresholds that were freed are reported in online supplemental table 4 of appendix A.

#### Relationship between the MHQoL dimensions and VAS score

Figure 1 presents the multilevel regression analysis examining the influence of MHQoL dimensions on the MHQoL-VAS, with an adjusted R<sup>2</sup> of 0.641. MHQoL-VAS scores were strongly associated with Self-Image, Daily activities, Mood and Future dimensions. Higher levels generally corresponded to higher VAS coefficients, except for Independence and Relationships, where the second category ('I am dissatisfied ...') showed a lower coefficient than a more severe level ('I am very dissatisfied ...'). However, these coefficients were not statistically significant, as their CIs crossed zero and extended into both negative and positive values. No significant difference in MHQoL scores was observed between male and female participants, and age was also not significantly associated with MHQoL scores.

#### Correlation between the EQ-5D-5L and MHQoL

Spearman's r coefficients (table 3) revealed moderate correlations between the MHQoL and both the EQ-5D-5L LSS and preference-based utilities, with stronger correlations for utilities, except in Finland. As expected,

correlations with EQ-5D-5L LSS were negative, whereas those with utilities were positive. For the MHQoL VAS, moderate correlations were observed in Spain, the UK and the overall sample across both EQ-5D-5L LSS and utilities, while Finland showed weaker correlations.

#### Relationship between the EQ-5D dimensions and MHQoL score

The results of the multilevel regression model to examine the influence of the EQ-5D-5L dimensions on the MHQoL score are presented in figure 2, with an adjusted R<sup>2</sup> of 0.565. As might be expected, the Anxiety/Depression dimension primarily influenced the MHQoL score. Additionally, some levels of the Pain/Discomfort, Usual Activities and Self-care also had a significant impact on the MHQoL score. Neither gender (male vs female) nor age showed a significant association with the MHQoL score.

#### Convergent validity

Correlations of the EQ-5D-5L and MHQoL with the disease-specific mental health questionnaires

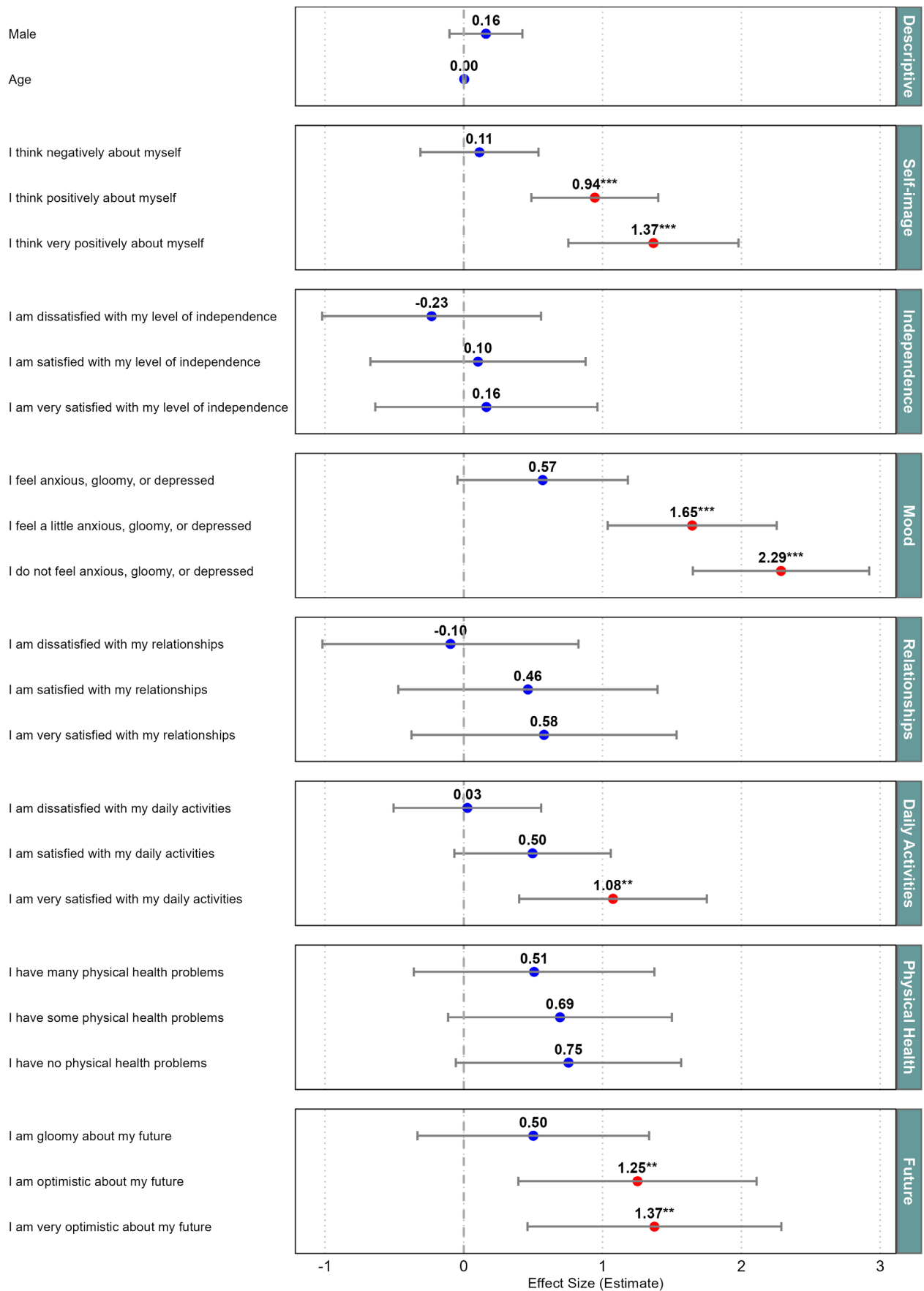
Correlation strengths varied across metrics and regions (table 4). Overall, the MHQoL showed stronger correlations than the EQ-5D-5L in all countries, except for the insomnia questionnaire (ISI) in Spain. The depression (PHQ-9) and well-being (WHO-5) questionnaires

**Table 2** Measurement invariance test

Model	CFI	RMSEA (90% CI)	SRMR	Model comparison	$\Delta\chi^2$ ( $\Delta$ df)	$\Delta$ CFI	$\Delta$ RMSEA	$\Delta$ SRMR
Model 1: Configural invariance	0.997	0.036 (0.000 to 0.064)	0.058	–	–	–	–	–
Model 2: Metric invariance	0.993	0.047 (0.017 to 0.070)	0.070	Model 1	17.347 (12)	–0.004	0.011	0.012
Model 3: Scalar invariance	0.961	0.095 (0.079 to 0.111)	0.065	Model 2	153.76* (20)	–0.032	0.048	–0.005
Model 3a: Partial scalar invariance	0.993	0.033 (0.013 to 0.067)	0.063	Model 2	3.251 (4)	0.000	–0.024	–0.070

\*p<0.001.

CFI, Comparative Fit Index; RMSEA, root mean square error of approximation; SRMR, standardised root mean square residual.



**Figure 1** Multilevel regression analysis of the effects of the MHQoL dimensions on the MHQoL VAS score, in which \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . MHQoL, Mental Health Quality of Life; VAS, Visual Analogue Scale.

**Table 3** Spearman's rank correlation coefficients expressed in Spearman's  $r$  for Finland, Spain, the UK and overall to present the correlation between the MHQoL score and MHQoL VAS with the EQ-5D-5L score and EQ-5D-5L utility

Spearman's $r$ <sup>1</sup>		
Country	EQ-5D-5L LSS	EQ-5D-5L Utility
Finland		
MHQoL	-0.565*	0.510*
MHQoL VAS	-0.332*	0.270
Spain		
MHQoL	-0.628*	0.687*
MHQoL VAS	-0.464*	0.528*
UK		
MHQoL	-0.568*	0.617*
MHQoL VAS	-0.496*	0.528*
Overall		
MHQoL	-0.585*	0.617*
MHQoL VAS	-0.450*	0.554*

\*p value<0.001  
EQ-5D-5L, EuroQol 5-Dimension 5-Level questionnaire; LSS, Level Sum Score; MHQoL, Mental Health Quality of Life; VAS, Visual Analogue Scale.

exhibited strong correlations with the MHQoL, while anxiety (GAD-7), insomnia (ISI), stress (PSS-4) and work stress (PRS) showed moderate correlations. For the EQ-5D-5L, the PHQ-9, GAD-7, ISI and PSS-4 were moderately correlated, while ISI and PRS showed weak correlations. In the UK, MHQoL correlations with ISI and PRS were weak, whereas EQ-5D-5L correlations followed the overall trend. In Finland, PSS-4 correlated strongly with MHQoL, but most EQ-5D-5L correlations with mental health questionnaires were weak, except for PHQ-9 (moderate). In Spain, all questionnaires except PRS showed moderate correlations with EQ-5D-5L, while MHQoL followed the overall correlation pattern.

### The MHQoL toolbox

Both the MHQoL R-package and the accompanying R-Shiny application are publicly available on CRAN under the names MHQoL<sup>47</sup> and MHQoL-Toolbox (<https://bit.ly/MHQoL-Toolbox>), respectively, indicating that it has successfully met rigorous quality standards, including comprehensive functionality testing, compliance with documentation requirements and proper dependency management. As a result, the package and tool are stable, well documented and openly accessible.

Independent verification by two authors (SP and FWT) confirmed that the package and tool consistently produced identical scores to manual MHQoL calculations, ensuring accuracy. Additionally, both calculation and result presentation were completed in under one second, significantly enhancing computation efficiency.

Because the code is publicly accessible on GitHub, a continuous feedback mechanism is established, allowing users to provide input and contribute to ongoing improvements in the user experience and addition of value sets.

## DISCUSSION

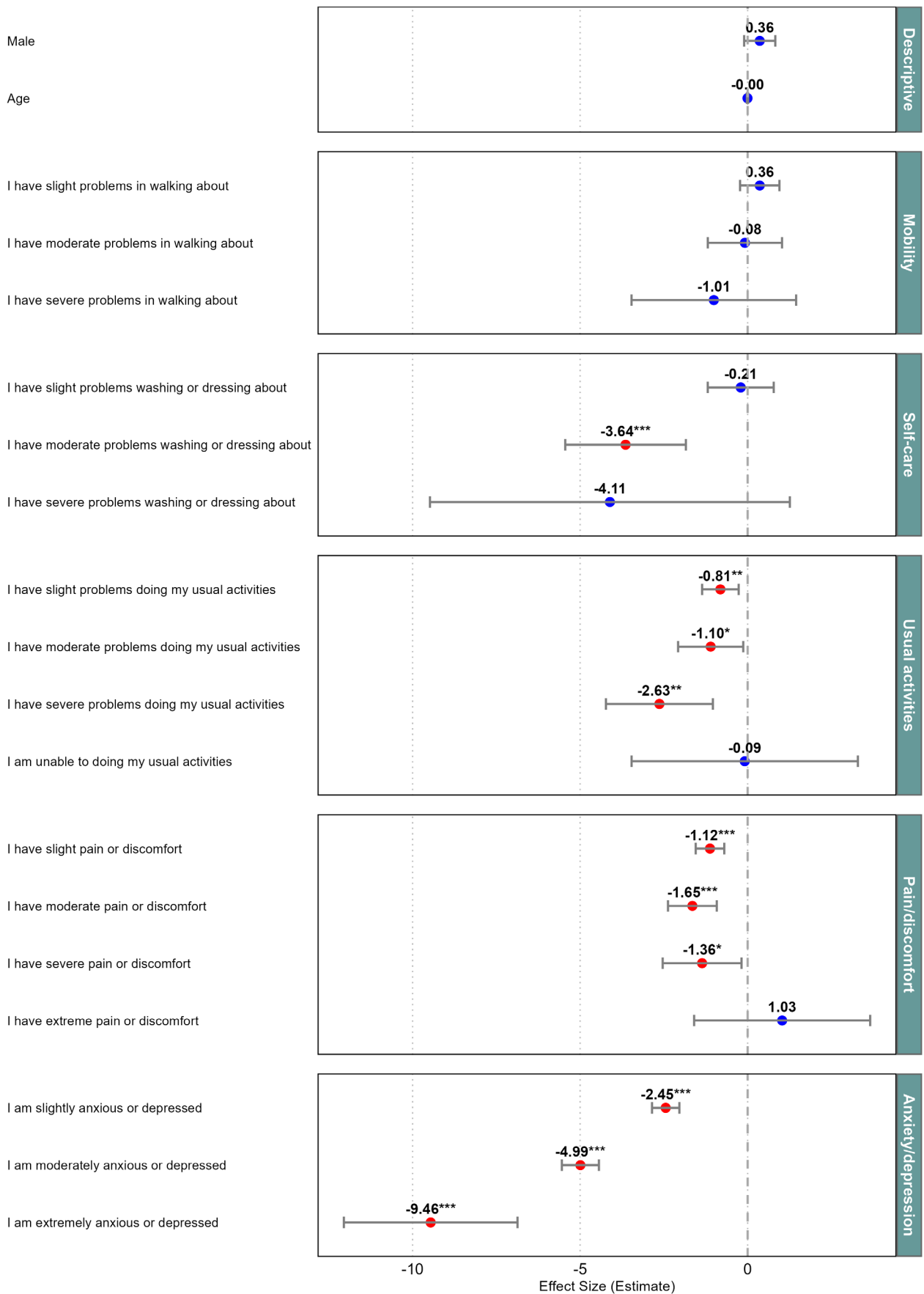
### Main results

Our findings indicate a strong and consistent relationship between the MHQoL, MHQoL- VAS and other mental health questionnaires across all participating countries. This supports the robustness of the MHQoL as a measure of mental health QoL. While statistical tests such as Cronbach's Alpha, McDonald's Omega, invariance testing and Spearman's rank correlation confirm these relationships, the primary implication is the reliability and validity of the MHQoL in a working population across different countries and contexts. This study established a solid foundation for the development of a standardised toolbox (R package and Shiny tool) enabling efficient MHQoL calculation, transformation and presentation, thereby streamlining the assessment process and enhancing its ease of use.

More specifically, the MHQoL demonstrated strong correlations not only among its own dimensions but also when compared with other questionnaires and between countries, underscoring its robust internal consistency and construct validity. In a comparative analysis with the EQ-5D-5L and disease-specific mental health questionnaires, the MHQoL consistently showed higher correlations with the disease-specific questionnaires across nearly all measures, although somewhat weaker associations between the MHQoL and other mental health instruments were observed in the Finnish sample. These differences may partly reflect the absence of a country-specific EQ-5D-5L value set for Finland, which required the use of the Danish value set as a proxy and may have influenced the estimated utility-based correlations.

Despite these overall strong associations, not all dimensions appear to be of equal importance. The multilevel analysis highlighted that the dimensions of Self-image, Mood, Daily activities and Future particularly impact the MHQoL VAS and thereby overall mental health QoL. From a broader perspective, it becomes apparent that not only mental health aspects correlate with QoL, but physical aspects do so as well. Specifically, dimensions such as Pain/Discomfort, Usual Activities and Self-care showed significant correlations.

compared with the psychometric studies investigating the properties of the MHQoL in the broader population and other countries, such as the study by Enzing *et al.*,<sup>20</sup> the MHQoL continues to demonstrate similarly strong psychometric properties, both for the internal consistency and construct validity. Although only the influence of the Anxiety/Depression dimension of the EQ-5D-5L on the MHQoL was measured, recommendations were made to further analyse the relationship between the MHQoL and EQ-5D-5L.<sup>20</sup>



**Figure 2** Multilevel regression analysis of the effects of the EQ-5D dimensions on the MHQoL score, in which \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . EQ-5D, Five-Dimension EuroQol Questionnaire; MHQoL, Mental Health Quality of Life.

**Table 4** Spearman's rank correlation coefficients expressed in Spearman's  $r$  for Finland, Spain, the UK and overall to present the correlation between the MHQoL and EQ-5D-5L with the disease-specific mental health questionnaires

Spearman's $r$ <sup>1</sup>		
Country	MHQoL	EQ-5D-5L
Finland		
PHQ-9	-0.790*	-0.426*
GAD-7	-0.570*	-0.267
ISI	-0.458*	-0.230
PSS-4	-0.719*	-0.275
PRS	-0.502*	-0.385*
WHO-5	0.766*	0.339*
Spain		
PHQ-9	-0.735*	-0.633*
GAD-7	-0.632*	-0.622*
ISI	-0.441*	-0.488*
PSS-4	-0.697*	-0.530*
PRS	-0.530*	-0.363*
WHO-5	0.790*	0.670*
UK		
PHQ-9	-0.727*	-0.586*
GAD-7	-0.607*	-0.446*
ISI	-0.348*	-0.326*
PSS-4	-0.698*	-0.454*
PRS	-0.330*	-0.237*
WHO-5	0.778*	0.573*
Overall		
PHQ-9	-0.750*	-0.585*
GAD-7	-0.605*	-0.491*
ISI	-0.395*	-0.342*
PSS-4	-0.698*	-0.479*
PRS	-0.397*	-0.278*
WHO-5	0.780*	0.589*

\* $p < 0.001$ .

EQ-5D-5L, EuroQol 5-Dimension 5-Level questionnaire; GAD-7, Generalised Anxiety Disorder-7; ISI, Insomnia Severity Index; MHQoL, Mental Health Quality of Life; PHQ-9, Patient Health Questionnaire-9; PRS, Psychosocial Risk Scale; PSS-4, Perceived Stress Scale-4; WHO-5, World Health Organization-Five Well-Being Index.

Our study demonstrated that the MHQoL appears to have an enhanced sensitivity in assessing mental health QoL compared with the EQ-5D-5L. However, outcomes of both the MHQoL and EQ-5D-5L seem to be interdependent, suggesting that the dimensions of these questionnaires influence each other's outcomes. This implies that while the MHQoL offers added value in capturing

mental health-specific dimensions, it may be most informative when used alongside generic instruments like the EQ-5D-5L. Together, these tools can provide a more comprehensive assessment of both general and mental health-related QoL, which is particularly relevant for economic evaluations and decision-making in health policy and workplace interventions.

### Strengths and limitations

A key strength of this study was the inclusion of a multi-country dataset, enabling cross-national comparisons and accounting for cultural differences in work environments. Additionally, reliability and validity were assessed using multiple approaches. Unlike prior studies focused solely on the Anxiety/Depression dimension of the EQ-5D-5L, this study conducted additional analyses to explore broader correlations. Furthermore, the development of a tool streamlined the transformation, calculation and presentation of the MHQoL results, enhancing usability and reproducibility in this and future studies. Importantly, this study was among the first to evaluate the MHQoL in a working population, providing critical insights for economic evaluations in occupational health research. Finally, the combination of psychometric validation with applied tool development supports both theoretical and practical advancement in the measurement of mental health-related QoL.

However, several limitations should be acknowledged. First, the sample included very few blue-collar employees, was predominantly female and highly educated, and largely consisted of public sector workers from three sectors that varied between countries. This may limit the generalisability of findings to the broader population. In addition, the voluntary nature of participation may have introduced selection bias, as employees with a greater interest in mental health may have been more likely to participate, which may be reflected in the observed sample. Second, most data originated from the UK, making the overall results less representative of Spain and Finland. Additionally, the absence of a country-specific value set for Finland's EQ-5D-5L preference-based utilities required the use of a proxy, potentially affecting the accuracy of correlations. Lastly, due to the relatively small sample size, certain EQ-5D-5L dimensions had low or no responses, leading to coefficients with high uncertainty or, in some cases, the inability to estimate coefficients. Furthermore, the analyses were based only on cross-sectional baseline data. Although the broader trial included follow-up measurements, these contained substantial missing data, which prevented reliable assessment of test-retest reliability. As a result, the present study could not evaluate temporal stability or responsiveness, and construct validation was limited to static cross-sectional patterns. In addition, the analyses relied primarily on self-reported measures (eg, MHQoL, MHQoL VAS and EQ-5D-5L). As a result, the observed associations may partly reflect common method variance, as responses were obtained using similar measurement approaches at the same time



point. This may have contributed to stronger correlations between measures and should be considered when interpreting the results. Finally, a limitation of this study is that, to date, a value set for the MHQoL is only available for the Netherlands,<sup>44</sup> which restricts its immediate use in economic evaluations across other countries. However, the developed tool is designed to be adaptable and will support the integration of additional country-specific value sets as they become available, facilitating future international applications.

### Recommendations and implications

The findings from this study, along with previous research,<sup>19 20</sup> demonstrate that the MHQoL is a valid and reliable questionnaire for measuring mental health QoL both in the general and working populations. Since all validation studies with MHQoL have been performed in Western Europe, extension to other regions would give more insight into the global applicability of this instrument. Furthermore, research in other groups of workers such as those in white-collar occupations, a higher proportion of men and individuals employed in business environments would provide a more comprehensive understanding of the working population. Moreover, gathering more disease-specific data is crucial. To achieve this, it is essential to incorporate the MHQoL in studies of mental health problems wherever possible.

In addition, this study provides preliminary evidence suggesting that the MHQoL may better capture the domains of mental health issues than the EQ-5D-5L. Further research is necessary to fully confirm this, both in the general and in the working population. If further evidence supports the positive outcomes observed, the MHQoL should be utilised to measure QoL in mental health issues alongside the EQ-5D-5L. This is because the questionnaires appear to influence each other's outcomes. Consequently, the relationship between the two questionnaires can be better examined, and a more comprehensive picture of the QoL can be depicted. However, for the application of the MHQoL in economic evaluations, data on the utilities of various dimensions are required from multiple countries; as of now, such data is available only for the Netherlands.<sup>44</sup>

### CONCLUSION

This study demonstrated that the MHQoL was a reliable and valid instrument for assessing mental health-related QoL in working populations across Finland, Spain and the United Kingdom. By focusing on employed individuals, this study also contributed to the growing field of economic evaluations targeting workplace mental health interventions, a domain of increasing societal and policy relevance due to the substantial burden of mental health problems on productivity, absenteeism and healthcare costs.

The results suggested that the MHQoL showed enhanced sensitivity to mental health domains compared

with general preference-based instruments such as the EQ-5D-5L and SF-36. Importantly, the MHQoL added value when used alongside these instruments in economic evaluations, by capturing aspects of QoL more directly related to mental well-being. Furthermore, a standardised and accessible toolbox, including an open-source R package and Shiny application, was developed to support the consistent scoring, transformation and reporting of MHQoL data, promoting broader application in both research and practice.

The use of a sensitive and targeted instrument like the MHQoL can help ensure that future evaluations more accurately capture the impact of mental health interventions in occupational settings.

### Author affiliations

<sup>1</sup>Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, Netherlands

<sup>2</sup>Department of Health Technology Assessment, Erasmus Universiteit Rotterdam Erasmus School of Health Policy and Management, Rotterdam, Netherlands

<sup>3</sup>Erasmus Universiteit Rotterdam Erasmus School of Health Policy and Management, Rotterdam, Netherlands

<sup>4</sup>Länsirannikon Työterveys, Turku, Finland

<sup>5</sup>Unit of Public Health, University of Turku, Turku, Finland

<sup>6</sup>Universitat de Barcelona, Barcelona, Spain

<sup>7</sup>CIBERSAM, Madrid, Spain

<sup>8</sup>Autonomous University of Barcelona, Barcelona, Spain

<sup>9</sup>Department of Health Sciences, York Biomedical Research Institute, University of York, Heslington, York, UK

<sup>10</sup>Hull York Medical School, University of York, York, UK

**Acknowledgements** The authors would like to thank the EMPOWER consortium for their support and contributions to this research

**Contributors** SBP, FWT and LH-VR all contributed to the conception and design of this study. All authors were involved in the writing process. SBP conducted the statistical analyses and created the MHQoL Toolbox with FWT and prepared the initial draft with the assistance of FWT, MDM and LH-VR. The remaining authors contributed to the final version of the manuscript. All authors read and approved the final manuscript. SBP is the guarantor of this study and accepts full responsibility for the work and/or the conduct of the study, had access to the data and controlled the decision to publish.

**Funding** The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 848180 and the National Health and Medical Research Council of Australia (APP1195937). B.O. is supported by the Miguel Servet (CP20/00040) contract, funded by the Instituto de Salud Carlos III and co-funded by the European Union (ERDF/ESF, "Investing in your future").

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Consent obtained directly from patient(s).

**Ethics approval** The study protocols have been approved by the ethics committees of Fundació Sant Joan de Déu (PIC-39-20), Turku University Hospital (PIC-993966082), University of York (HSRGC250321) and Institute of Occupational Medicine, University of Lodz (9/2020). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content

includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

Stijn B Peeters <https://orcid.org/0009-0004-3684-3584>

Marleen De Mul <https://orcid.org/0000-0003-3447-5920>

Marjo Sinokki <https://orcid.org/0000-0001-9546-4905>

Christina Maria Van Der Feltz-Cornelis <https://orcid.org/0000-0001-6925-8956>

#### REFERENCES

- Global health data exchange (GHDX). Institute of Health Metrics and Evaluation.
- Mental health and covid-19: early evidence of the pandemic's impact*. World Health Organization (WHO), 2022.
- Shaholli D, Mantione G, Torre G. Work-related stress costs assessment: a narrative review. *Clin Ter* 2023;456–60.
- EU Compass for Action on Mental Health and Well-Being. European Institute for Gender Equality; 2016. Available: <https://policycommons.net/artifacts/1943653/eu-compass-for-action-on-mental-health-and-well-being/2695422/> [Accessed 15 Aug 2023].
- Bloom DE, Califano E, Jané-Llopis E, et al. *The global economic burden of non-communicable diseases*. World Economic Forum, 2012.
- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. 2023.
- de Oliveira C, Saka M, Bone L, et al. The Role of Mental Health on Workplace Productivity: A Critical Review of the Literature. *Appl Health Econ Health Policy* 2023;21:167–93.
- Duchaine CS, Aubé K, Gilbert-Ouimet M, et al. Psychosocial Stressors at Work and the Risk of Sickness Absence Due to a Diagnosed Mental Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry* 2020;77:842–51.
- Bryan ML, Bryce AM, Roberts J. The effect of mental and physical health problems on sickness absence. *Eur J Health Econ* 2021;22:1519–33.
- Parsaei R, Roohafza H, Feizi A, et al. How Different Stressors Affect Quality of Life: An Application of Multilevel Latent Class Analysis on a Large Sample of Industrial Employees. *Risk Manag Healthc Policy* 2020;13:1261–70.
- Rugulies R, Aust B, Greiner BA, et al. Work-related causes of mental health conditions and interventions for their improvement in workplaces. *The Lancet* 2023;402:1368–81.
- Babapour AR, Gahassab-Mozaffari N, Fathnezhad-Kazemi A. Nurses' job stress and its impact on quality of life and caring behaviors: a cross-sectional study. *BMC Nurs* 2022;21:75.
- Bergman E, Löytyniemi E, Myllyntausta S, et al. Factors associated with quality of life and work ability among Finnish municipal employees: a cross-sectional study. *BMJ Open* 2020;10:e035544.
- Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). *Med Care* 1992;30:473–83.
- Balestroni G, Bertolotti G. EuroQoL-5D (EQ-5D): an instrument for measuring quality of life. *Monaldi Arch Chest Dis* 2015;78.
- Brazier J, Connell J, Papaioannou D, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol Assess (Rockv)* 2014;18.
- Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: Are we asking the right questions? *Social Science & Medicine* 2014;120:12–20.
- Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. *Value Health* 2011;14:907–20.
- van Krugten FCW, Busschbach JJV, Versteegh MM, et al. The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Qual Life Res* 2022;31:633–43.
- Enzing JJ, van Krugten FCW, Sabat I, et al. Psychometric evaluation of the Mental Health Quality of Life (MHQoL) instrument in seven European countries. *Health Qual Life Outcomes* 2022;20:129.
- Minds O. Fitter Jobs: From Awareness to Change in Integrated Mental Health. *Skills Work Pol OECD* 2021.
- Olaya B, Van der Feltz-Cornelis CM, Hakkaart-van Roijen L, et al. Study protocol of EMPOWER: A cluster randomized trial of a multimodal eHealth intervention for promoting mental health in the workplace following a stepped wedge trial design. *Digit Health* 2022;8:20552076221131145.
- Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18.
- Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- Methods for analysing and reporting EQ-5D data*. Springer Nature, 2021.
- Ramos-Goñi JM, Craig BM, Oppe M, et al. Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach. *Value Health* 2018;21:596–604.
- Devlin NJ, Shah KK, Feng Y, et al. Valuing health-related quality of life: An EQ-5D-5L value set for England. *Health Econ* 2018;27:7–22.
- Jensen CE, Sørensen SS, Gudex C, et al. The Danish EQ-5D-5L Value Set: A Hybrid Model Using cTTO and DCE Data. *Appl Health Econ Health Policy* 2021;19:579–91.
- Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–13.
- Spitzer RL, Kroenke K, Williams JBW, et al. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med* 2006;166.
- Morin CM. *Insomnia: psychological assessment and management*. Guilford Press, 1993:238.
- Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297–307.
- Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:385–385–96.
- Najder A, Merecz-Kot D, Wójcik A. Relationships between occupational functioning and stress among radio journalists – Assessment by means of the psychosocial risk scale. *Int J Occup Med Environ Health* 2016;29:85–100.
- Topp CW, Østergaard SD, Søndergaard S, et al. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom* 2015;84:167–76.
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Medical Education* 2011;2:53–5.
- Putnick DL, Bornstein MH. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Dev Rev* 2016;41:71–90.
- Bowen NK, Masa RD. Conducting Measurement Invariance Tests with Ordinal Data: A Guide for Social Work Researchers. *J Soc Social Work Res* 2015;6:229–49.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model Multidiscipl J* 1999;6:1–55.
- Hauke J, Kossowski T. Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *Quaestiones Geograph* 2011;30:87–93.
- Casella G, Berger RL. *Statistical inference*. 2nd edn. CENGAGE Learning, 2002.
- Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analgesia* 2018;126:1763–8.
- R Core Team. R: a language and environment for statistical computing. 2021. Available: <https://www.R-project.org/>
- KrugtenFCWJonker MF, Himmler SFW, et al. Estimating a Preference-Based Value Set for the Mental Health Quality of Life Questionnaire (MHQoL). *Med Decis Making* 2024;44:64–75.
- Janssen B, Szende A. Population norms for the EQ-5D. In: Szende A, Janssen B, Cabases J, eds. *Self-reported population health: an international perspective based on EQ-5D*. Springer Netherlands, 2014: 19–30.
- van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. *Eur J Devel Psychol* 2012;9:486–92.
- Peeters S, Thielen F. MHQoL: Mental Health Quality of Life ToolKit. 2025.