

The WHO-5 Well-Being Index: A Systematic Review of the Literature

Christian Winther Topp^a Søren Dinesen Østergaard^{b, c} Susan Søndergaard^a
Per Bech^a

^aPsychiatric Research Unit, Psychiatric Centre North Zealand, Copenhagen University Hospital, Hillerød,

^bDepartment of Clinical Medicine, Aarhus University Hospital, Aarhus University, Aarhus, and ^cDepartment P – Research, Aarhus University Hospital, Risskov, Risskov, Denmark

Key Words

WHO-5 Well-Being Index · Systematic review of literature · Depression screening tool · Sensitivity and specificity · Clinimetrics

Abstract

Background: The 5-item World Health Organization Well-Being Index (WHO-5) is among the most widely used questionnaires assessing subjective psychological well-being. Since its first publication in 1998, the WHO-5 has been translated into more than 30 languages and has been used in research studies all over the world. We now provide a systematic review of the literature on the WHO-5. **Methods:** We conducted a systematic search for literature on the WHO-5 in PubMed and PsycINFO in accordance with the PRISMA guidelines. In our review of the identified articles, we focused particularly on the following aspects: (1) the clinimetric validity of the WHO-5; (2) the responsiveness/sensitivity of the WHO-5 in controlled clinical trials; (3) the potential of the WHO-5 as a screening tool for depression, and (4) the applicability of the WHO-5 across study fields. **Results:** A total of 213 articles met the predefined criteria for inclusion in the review. The review demonstrated that the WHO-5 has high clinimetric validity, can be used as an outcome measure balancing the wanted and unwanted effects of treatments, is a

sensitive and specific screening tool for depression and its applicability across study fields is very high. **Conclusions:** The WHO-5 is a short questionnaire consisting of 5 simple and non-invasive questions, which tap into the subjective well-being of the respondents. The scale has adequate validity both as a screening tool for depression and as an outcome measure in clinical trials and has been applied successfully across a wide range of study fields.

© 2015 S. Karger AG, Basel

Introduction

In his monograph on clinimetrics, Feinstein [1] used the term ‘improvement after treatment’ to describe the patient’s own assessment of change in his or her well-being during treatment. This is a very subjective index, which no biological marker can capture [1–3]. When reviewing 75 scientific articles covering more than 100 different scales or questionnaires, Gill and Feinstein [4] demonstrated that a clinimetric definition of subjective well-being was lacking. They therefore advocated for the development of short global rating scales of subjective well-being, which would reflect a single dimension with high clinical face validity. Another important issue in relation to the measurement of subjective well-being was

| The WHO-5 questionnaire | | | | | | |
|---|-----------------|------------------|-------------------------|-------------------------|------------------|------------|
| Instructions: | | | | | | |
| Please indicate for each of the 5 statements which is closest to how you have been feeling over the past 2 weeks. | | | | | | |
| Over the past 2 weeks... | All of the time | Most of the time | More than half the time | Less than half the time | Some of the time | At no time |
| 1 ... I have felt cheerful and in good spirits | 5 | 4 | 3 | 2 | 1 | 0 |
| 2 ... I have felt calm and relaxed | 5 | 4 | 3 | 2 | 1 | 0 |
| 3 ... I have felt active and vigorous | 5 | 4 | 3 | 2 | 1 | 0 |
| 4 ... I woke up feeling fresh and rested | 5 | 4 | 3 | 2 | 1 | 0 |
| 5 ... my daily life has been filled with things that interest me | 5 | 4 | 3 | 2 | 1 | 0 |

Scoring principle: The raw score ranging from 0 to 25 is multiplied by 4 to give the final score from 0 representing the worst imaginable well-being to 100 representing the best imaginable well-being.

Fig. 1. The WHO-5 Well-Being Scale. Instructions and scoring principle.

raised by Ware [5], who suggested that the rating scales dedicated to this purpose should be disease anonymous (generic) because such scales provide information regarding the overall effect (balancing wanted clinical effects against unwanted adverse effects) of a clinical intervention. Furthermore, a generic scale enables a comparison with mean values from the general population (which can be used as a criterion of remission) or with mean values from other clinical populations irrespective of the disease entity or condition under examination.

The 5-item World Health Organization Well-Being Index (WHO-5) is a short and generic global rating scale measuring subjective well-being. The WHO-5 was derived from the WHO-10 [6], which in turn was derived from a 28-item rating scale [7] used in a WHO multicentre study in 8 different European countries [8]. The 10 items making up the WHO-10 were selected from among these 28 items on the basis of a non-parametric item response theory analysis [9], which identified the 10 most valid items from the original 28-item scale [7]. The items for the 28-item scale were selected from the Zung scales for depression, distress and anxiety as well as from the General Health Questionnaire and the Psychological General Well-Being Scale [10]. Therefore, both the 28-item scale and the WHO-10 include items phrased negatively to reflect symptoms of distress ('Feeling downhearted and blue') and items phrased positively, reflecting well-being ('Waking up feeling fresh and rested'). Because the WHO considers positive well-being to be another term for mental health [11], the WHO-5 only contains positively phrased items [12]. The WHO-5 items (fig. 1) are: (1) 'I have felt cheerful and in good spirits', (2) 'I have felt calm and relaxed', (3) 'I have

felt active and vigorous', (4) 'I woke up feeling fresh and rested' and (5) 'My daily life has been filled with things that interest me'. The respondent is asked to rate how well each of the 5 statements applies to him or her when considering the last 14 days. Each of the 5 items is scored from 5 (all of the time) to 0 (none of the time). The raw score therefore theoretically ranges from 0 (absence of well-being) to 25 (maximal well-being). Because scales measuring health-related quality of life are conventionally translated to a percentage scale from 0 (absent) to 100 (maximal), it is recommended to multiply the raw score by 4 (fig. 1). Notably, the layout of the WHO-5 follows that of the Major Depression Inventory which measures the WHO/ICD-10 symptoms of depression [13]. This goes both for the Likert scaling of each item from 0 to 5 and for the period of time considered (the past 2 weeks) [14, 15].

The WHO-5 was originally presented at a WHO meeting in Stockholm in February 1998 as part of a project on the measurement of well-being in primary health care patients [16]. Subsequent to this project, the WHO Regional Office in Europe initiated translations of the original English version of the WHO-5 into a number of other languages. At present, the WHO-5 has been translated into over 30 languages and has been used in research projects all over the world. The objective of this study was to provide a systematic review of the extensive body of literature on the WHO-5, with particular emphasis on the following aspects: (1) the clinimetric validity of the WHO-5; (2) the responsiveness/sensitivity of the WHO-5 in controlled clinical trials; (3) the potential of the WHO-5 as a screening tool for depression, and (4) the applicability of the WHO-5 across study fields.

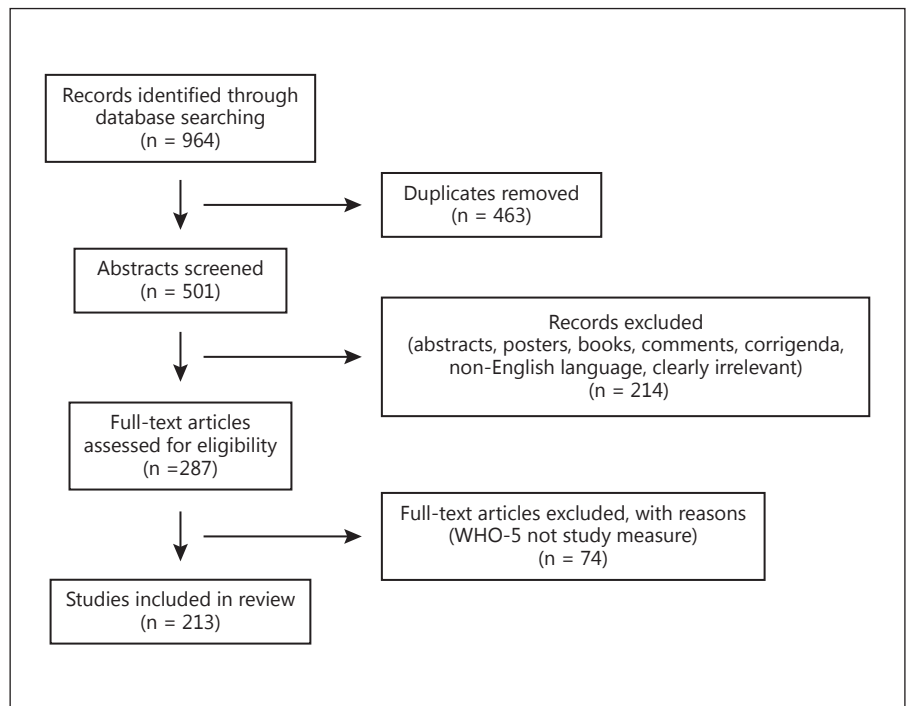


Fig. 2. PRISMA flow diagram showing the flow of information through the review: the number of records identified, included and excluded, and the reasons for the exclusions.

Methods

Search Strategy

A search for 'WHO (Five)' OR 'WHO-5' OR 'WHO-five' OR 'WHO well-being' OR 'WHO well being' OR 'WHO 5' OR 'WHO five' OR 'World health organization 5' or 'World health organization five' was carried out in PubMed. The same search was then carried out in PsycINFO. All abstracts published prior to or on March 31, 2014 were considered. One author (S.S.) screened the search results to exclude books, conference abstracts/posters and papers that were clearly irrelevant as well as non-English articles. The remaining papers were evaluated by C.W.T. and S.S. and were included for full text review if they contained information on the use of the WHO-5 other than as a pre-study screening instrument. C.W.T., S.S. and P.B. then evaluated the full papers.

Results

The PRISMA flow chart [17] (fig. 2) shows the number of articles found and later kept or excluded in the different phases of the screening. The database search identified 964 titles, which were reduced to 501 after removal of duplicates. A total of 214 of these titles were either not full-text articles, were written in a non-English language or were obviously not on a WHO-5-related subject and therefore excluded. The remaining 287 full-text articles were screened, and 213 were found eligible for in-

clusion in the review. In online supplementary table 1 (for all online suppl. material, see www.karger.com/doi/10.1159/000376585), these 213 published studies are listed, and the main WHO-5 results are described. Besides the trials listed in online supplementary table 1, we consulted the European Quality of Life Survey 2012 [18] to obtain the European general population norm values on the WHO-5 (online suppl. table 2).

The Clinimetric Validity of the WHO-5

The most adequate clinimetric evaluation of the WHO-5 was performed by a panel of experts in the field of health-related quality of life [19]. This group evaluated 85 different questionnaires and found that 20 of these were 'acceptable'. In terms of clinimetric validity, the WHO-5 was listed at the top among the 20 scales since any major overlap with specific disease-related aspects and side effects of pharmacological treatment is absent on this scale. In other words, the WHO-5 is a pure generic scale for the measurement of general well-being [19].

The construct validity of a scale describes its properties as a coherent measure of a dimension of interest (in this case well-being). Construct validity is evaluated by determining whether each item on the scale contributes with unique information regarding the dimension. If this is the case, the scale covers the theoretical range from the

Table 1. The 6 controlled clinical trials in which the WHO-5 has been used as outcome measure

| Study | Intervention | Specific condition | Assessment occasion | Active intervention | Control intervention | p |
|-------|--|--------------------------|---------------------------------|-----------------------------|-----------------------------|-------|
| 1 | Prolonged-release melatonin versus placebo [26] | Psychiatry – other | Baseline End point | (n = 169) 64.0 69.2 | (n = 165) 62.0 66.4 | 0.05 |
| 2 | Mindfulness-based therapy versus wait list [28] | Oncology | Baseline End point | (n = 103) 52.2 60.3 | (n = 111) 50.1 50.7 | 0.01 |
| 3 | Intraperitoneal versus subcutaneous administration of insulin [29] | Endocrinology (diabetes) | Baseline End point | (n = 11) 48.4 67.6 | (n = 12) 43.7 45.0 | 0.05 |
| 4 | Paroxetine versus placebo [30] | Otolaryngology | Baseline End point | (n = 57) 25.2 25.1 | (n = 58) 24.6 25.4 | >0.05 |
| 5 | Desvenlafaxine versus placebo [31] | Depression | Baseline End point (8 weeks) | (n = 1,925) 23.6 50.8 | (n = 1,178) 23.6 44.0 | <0.01 |
| 6 | Wake therapy versus exercise [32] | Depression | Baseline End point (9 weeks) | (n = 37) 12.0 58.0 | (n = 38) 18.0 50.0 | <0.05 |

complete absence of well-being to the highest imaginable level of well-being. The WHO-5 has been analysed with the item response theory model formulated by Rasch [20] in both younger persons [21] and in elderly persons [22], confirming that the 5 items constitute a unidimensional scale, where each item adds unique information regarding the level of well-being. This implies that the individual item scores can be added to a ‘meaningful’ total score and that the range of scores from 0 to 100 covers the entire dimension from the complete absence of well-being to the highest imaginable level of well-being.

The mean WHO-5 score in the general population has been measured in different European countries (online suppl. table 2). Thus, when using the WHO-5 as an outcome measure in clinical trials, the ideal goal should be to reach the general population mean score. In Danish general population studies [23, 24], the mean WHO-5 score is 70.

The predictive validity of a rating scale is of great importance. The predictive validity of the WHO-5 has been investigated in a study in which patients with cardiac disease were followed over a period of 6 years [25]. Patients who scored <50 on the WHO-5 at baseline proved to have significantly higher mortality rates compared to those scoring ≥ 50 .

The Responsiveness/Sensitivity of the WHO-5 in Controlled Clinical Trials

Table 1 describes the 6 controlled clinical trials in which the WHO-5 has been used as an outcome measure. Wade et al. [26] conducted a placebo-controlled study of melatonin for the treatment of insomnia. At the end point, the participants in the group had obtained a level of 69. However, the difference at the end point between the active group and the control group was approximately 3 on the WHO-5, which is statistically significant but not clinically significant, since the threshold for a clinically relevant change is considered to be 10 points on the WHO-5 [27].

Hoffman et al. [28] tested the effect of mindfulness-based therapy versus a wait-list group among patients with breast cancer. The WHO-5 baseline score in each of the groups was approximately 50, which is indicative of reduced well-being (when WHO-5 is used for the screening of depression, a cut-off score of ≤ 50 is used; table 2). The difference between the effect of the mindfulness-based therapy and the control group was approximately 10 points on the WHO-5, i.e. just barely clinically significant, but the patients in the active group still had mean WHO-5 values below the general population norm at the end point [28].

Table 2. Diagnostic accuracy of WHO-5 using a cut-off score of ≤ 50

| Study [Ref] | Condition/topic | Index of validity | Sensitivity | Specificity |
|----------------------------|---------------------------|----------------------------|-------------|-------------|
| 1 [64] | Geriatrics | DSM-IV depression | 0.92 | 0.79 |
| 2 [65] | Endocrinology (diabetes) | SCID DSM-IV depression | 1.00 | 0.78 |
| 3 [66] | Clinical psychometrics | SCID DSM-IV depression | 0.77 | 0.89 |
| 4 [67] | Neurology | SCID DSM-IV depression | 0.78 | 0.67 |
| 5 [68] | Psychiatry – other | DSM-IV depression | 0.96 | 0.76 |
| 6 [69] | Clinical psychometrics | MINI DSM-IV depression | 0.89 | 0.65 |
| 7 [46] | Neurology | MINI DSM-IV depression | 0.88 | 0.74 |
| 8 [70] | Geriatrics | DSM-IV depression | 0.78 | 0.83 |
| 9 [33] | Clinical psychometrics | DSM-IV depression | 0.94 | 0.83 |
| 10 [71] | <i>Paediatrics</i> | | | |
| | Children (9–12 years) | DSM-IV depression | 0.75 | 0.92 |
| | Adolescents (13–16 years) | DSM-IV depression | 0.74 | 0.89 |
| 11 [72] | Geriatrics | CIDI ICD-10 depression | 1.00 | 0.80 |
| 12 [73] | Depression | CIDI ICD-10 depression | 0.93 | 0.64 |
| 13 [74] | Clinical psychometrics | CIDI ICD-10 depression | 0.90 | 0.63 |
| 14 [75] | Endocrinology (diabetes) | CES-D ≥ 16 depression | 0.89 | 0.87 |
| 15 [76] | Endocrinology (diabetes) | CES-D ≥ 16 depression | 0.57 | 0.83 |
| 16 [77] | Endocrinology (diabetes) | PHQ9 ≥ 10 depression | 0.79 | 0.88 |
| 17 [78] | Geriatrics | MDI DSM-IV depression | 0.97 | 0.86 |
| 18 [79] | Geriatrics | CIDI suicidal | 0.87 | 0.75 |
| All 18 studies (n = 5,823) | | | 0.86 | 0.81 |

In study 9 [33], the cut-off score was ≤ 28 .

Logtenberg et al. [29] compared the effects of two different forms of insulin administration in patients with type 1 diabetes. At baseline, the participants had a WHO-5 mean score just below 50, and those in the active group (intraperitoneal administration) improved their WHO-5 scores, reaching the general population norm, in contrast to the control group (subcutaneous administration).

Robinson et al. [30] tested the effect of paroxetine in comparison to placebo among individuals with tinnitus. The baseline WHO-5 mean score was below the cut-off score for clinical depression (≤ 28). According to the WHO-5 results, neither paroxetine nor placebo had any effect on the well-being of those affected by tinnitus [30].

The WHO-5 has also been used as an outcome measure in controlled clinical trials involving patients with major depression. In these trials, the WHO-5 can be viewed as a measure which taps into the balance between the desired clinical effects and the unwanted side effects.

The study by Guico-Pabia et al. [31] compared desvenlafaxine with placebo by pooling the results from 8 different trials. The baseline score is representative for major depression in the primary care setting (mean score ≤ 28). The difference between desvenlafaxine and placebo after 8 weeks was statistically significant but not clinically significant. Also, the desvenlafaxine group only reached a WHO-5 mean of 50, i.e. it still has not recovered if we use the general population norm as the goal of treatment [31].

Another study of depression assessed the effect of wake therapy and exercise among patients with treatment-resistant depression, i.e. those who had not responded to at least two different antidepressants in their current major depressive episode [32]. During the 9-week trial, the increase in the WHO-5 score in the wake therapy group was statistically superior to that of the exercise group. However, the end point WHO-5 mean score in both groups was still far below that of the general population norm [32].

Table 3. The applicability of the WHO-5 across study fields

| Fields or conditions | Publications, n |
|------------------------|-----------------|
| Endocrinology | 34 |
| Depression | 29 |
| Stress | 29 |
| Psychology | 22 |
| Clinical psychometrics | 21 |
| Psychiatry – other | 19 |
| Geriatrics | 16 |
| Neurology | 9 |
| Cardiology | 7 |
| Oncology | 6 |
| Obstetrics | 5 |
| Pain | 5 |
| Suicidology | 5 |
| Paediatrics | 2 |
| Gynaecology | 1 |
| Ophthalmology | 1 |
| Otolaryngology | 1 |
| Health economics | 1 |

The Potential of the WHO-5 as a Screening Tool for Depression

Table 2 describes the performance (sensitivity and specificity) of the WHO-5 when used as a screening tool for depression. Most of the studies used DSM-IV depression assessed by various structured interviews as their gold standard reference. In the first 8 studies listed in table 2, a cut-off score of ≤ 50 on the WHO-5 was used to assign a 'screening diagnosis' of depression. For these 8 studies, the mean sensitivity for DSM-IV depression was 0.87, and the mean specificity for DSM major depression was 0.76. For all 18 studies, the sensitivity of the WHO-5 was 0.86, and the specificity was 0.81 (table 2). In the study by Lowe et al. [33], the WHO-5 cut-off score was ≤ 28 , which more restrictively equals the level of well-being among patients with DSM-IV major depression. Despite this restrictive threshold, the WHO-5 had a sensitivity of 0.93 and a specificity of 0.83 in the detection of depression.

For a screening instrument such as the WHO-5, having sufficiently high sensitivity (i.e. a very high proportion of depressed individuals screen positive) is a key factor, whereas high specificity is less important. This is due to the fact that the second step of the diagnostic process, after an initial positive screening with the WHO-5, consists of a diagnostic interview performed by a trained clinician, during which 'false positives' (patients screening positive on the WHO-5 but not meeting criteria for depression) will be detected [16].

The Applicability of the WHO-5 across Study Fields

Table 3 lists the number of WHO-5 publications stratified by field of study. The scale has been used most extensively in endocrinology, which is explained by the fact that the WHO-5 was developed in a Pan-European study of patients with diabetes [6]. Within the field of diabetes, the most comprehensive study employing the WHO-5 is the multinational study of psychosocial issues in diabetes by Nicolucci et al. [34]. The participants were patients with diabetes from 17 countries. Using a cut-off score of ≤ 28 on the WHO-5, approximately 14% of the patients were screened as having a depression. As shown by Nicolucci et al. [34], depression may have a substantial negative impact on diabetic control.

Among the WHO-5 studies on depression, the work by Krieger et al. [35] is highly important because both the Hamilton Depression Scale (HAM-D) and the Hamilton Anxiety Scale (HAM-A) were included among the indices of validity. A score of 18 on the HAM-D, which indicates the level of depression at which treatment is needed (major depression), corresponded to a WHO-5 score of 20, and an HAM-D score of 13 (minor depression) corresponded to a WHO-5 score of 32 [35].

Suicidology is also among the fields where the WHO-5 has been applied. When health-related quality of life was accepted as an important outcome in clinical trials, it was found that psychological well-being was a major element of this dimension. As discussed by Andrews and Withey [36], psychological well-being can be considered as the sum of satisfactions that makes life worth living, constituting the opposite pole to psychological pain with suicidal thoughts [36]. In the WHO Multisite Intervention Study on Suicidal Behaviours (SUPRE-MISS), the WHO-5 was used to identify subjects who attempted suicide. The main finding of the SUPRE-MISS was that those attempting suicide generally had extremely low scores on the WHO-5 [37, 38]. Furthermore, Vijayakumar et al. [39] found that subjects with repeated suicide attempts scored lower on the WHO-5 than subjects with a first attempt (49.6 vs. 58.6, $p = 0.01$). Finally, Awata et al. [40] showed that subjects with suicidal ideation scored significantly lower on the WHO-5 than subjects without suicidal ideation (45.6 vs. 67.6, $p = 0.01$).

In other illnesses than depression within the psychiatric field, the WHO-5 has been used most extensively in relation to alcoholism and other substance use disorders. Hensing et al. [41] found a negative correlation between well-being and harmful alcohol use in females but not in males. Elholm et al. [42] focused on the alcohol withdrawal syndrome in outpatients treated with chlordiazepoxide

as anti-abstinence drug. Most individuals scored <50 on the WHO-5 at inclusion. After 2 weeks of therapy, the WHO-5 score increased significantly. Hoxmark et al. [43] found that the mean scores were approximately 38 on the WHO-5 at baseline and that 47% scored <50. There was a negative correlation between the severity of abuse and the WHO-5 scores.

Cardiovascular disease has previously been mentioned in the section on the clinimetric properties of the WHO-5 [25]. The results found by Birket-Smith et al. [25], i.e. that myocardial infarction patients with WHO-5 scores of >50 are well functioning, was supported by the study by Bergmann et al. [44] among patients who had survived myocardial infarction and were followed for >2 years. The fact that patients having survived myocardial infarction have rather high WHO-5 scores has been considered by Garnefski et al. [45] as an expression of a positive coping style. They referred to this effect as a form of 'post-traumatic growth'.

In the field of neurology, the WHO-5 has been found valid in screening for depression in patients with Parkinson's disease [46]. Furthermore, the WHO-5 has been used as an outcome scale in a placebo-controlled study on a new pain-reducing wound dressing in patients with venous leg ulcers, where the WHO-5 was used as an outcome scale [47]. The results showed a significant superiority of the active treatment over placebo. The applicability of WHO-5 in back pain disorder has also been evaluated as being acceptable by both Volinn et al. [48] and Vereckei et al. [49].

In the field of stress research, the WHO-5 has been used to assess a wide variety of aspects including coping strategies [50], well-being in occupational health settings [51], the association between workplace stress and well-being [52], the links between working condition and well-being [53] as well as the association between psychosocial conditions and well-being [54]. In the study by Gao et al. [52], it was found that approximately 35% of a total of 2,796 employees had low well-being (cut-off score on the WHO-5 of <50) and that low social capital at the workplace was associated with poor well-being. This finding was confirmed by Jung et al. [55]. Finally, in the study by Schutte et al. [54] it was reported that the prevalence of poor well-being was highest in the low education group.

Discussion

In accordance with the suggestion from Feinstein [1] and Ware [5], the WHO-5 was developed as a generic scale without any diagnostic specificity [6]. The WHO-5

was therefore recommended as the ultimate patient-related measure within the international WHO classification system for chronic medical conditions, integrating biological impairments, social disabilities and subjective handicaps [56]. In this systematic review of the WHO-5 as a generic well-being scale, we focused on four aspects: (1) the clinimetric validity of the WHO-5; (2) the responsiveness/sensitivity of the WHO-5 in controlled clinical trials; (3) the potential of the WHO-5 as a screening tool for depression, and (4) the applicability of the WHO-5 across study fields. It is our impression that the WHO-5 has performed well with regard to all these aspects.

The clinimetric validity of the WHO-5 has been evaluated in terms of construct validity (total score being a sufficient statistic). Item response theory analyses have shown that the WHO-5 covers the dimension of subjective well-being from 0 (worst imaginable well-being) to 100 (best imaginable well-being). In the comprehensive review by Hall et al. [19], the clinical validity of the WHO-5 was evaluated to be very high as the scale can be used irrespective of underlying illness (or lack of illness) and across many different settings.

When used as an outcome measure, the WHO-5 has been able to capture improvement in well-being caused by various pharmacological and non-pharmacological interventions in clinical trials across many different branches of medicine [28, 30, 32]. Furthermore, by comparing end point ratings on the WHO-5 to general population mean scores, remission rates can be calculated. When used as an outcome measure in clinical trials, the WHO-5 can be viewed as a measure which taps into the balance between the desired clinical effects and the unwanted side effects of a given intervention [20].

Feinstein [1] suggested that the validity of a clinimetric scale like the WHO-5 should be tested in terms of sensitivity and specificity, analogously to erythrocyte sedimentation rates as a diagnostic test in general medicine [57]. When summarizing all WHO-5 studies for the screening of depression, the weighted sensitivity was 0.86 and the specificity 0.81 (table 2), which is acceptable. Using a WHO-5 cut-off score of ≤ 50 is recommendable when screening for clinical depression.

In the late 1990s, the WHO-5 was introduced in the medical field of diabetes. Our review has shown that diabetes remains the condition in which the WHO-5 has been used most extensively. The WHO-5 studies in the field of diabetes indicate that a low WHO-5 score may have a substantial negative impact upon diabetic control, thereby validating the use of the WHO-5 within the international WHO classification system for chronic medical conditions

[56]. In this context, the biological impairment is the metabolic disturbance caused by diabetes, and the decreased well-being is an indication of problems in the diabetic care of the patient. This in contrast to some cardiovascular disorders or cancer disorders, where a low WHO-5 score might indicate a clinical depression where the impairment is depression-related, for example poststroke depression [58] or cancer depression [59]. Within diabetes, obesity and metabolic syndrome have been shown to be related to depression [60]. In a recent review on the metabolic syndrome, Bergmann et al. [61] found that scales focusing on stressor rather than on distress or low well-being had been used. Many stress-related studies have been performed using the WHO-5 to measure distress or poor well-being. Also, within the WHO SUPRE-MISS, the WHO-5 has been found to be highly applicable [38].

The WHO-5 has been used extensively worldwide. Online supplementary table 1 shows the diversity of its application across different regions: Africa (Algeria, South Africa), Asia (Bangladesh, China, India, Japan, South Korea, Sri Lanka, Taiwan, Thailand), Europe (Northern, Southern, Eastern, Western and Central Europe), the Americas (Canada, the US, Brazil, Mexico), the Middle East (Israel, Iran, Lebanon) and Oceania (Australia, New Zealand). This very successful dissemination of the scale is probably due to its straightforward language, which poses few translation problems and to the fact that the questions do not seem to transgress any cultural norms in the individual countries.

Recently, Ryff [62] has reviewed publications on the eudaemonic scales of well-being, which capture the core

aspects of what it means to be human, including existential and developmental factors. Based on this eudaemonic approach [11], Fava [63] developed his 'well-being therapy'. It should be stated that the WHO-5 is a clinimetric outcome scale at the description level, parallel to symptom-related scales and side effect scales. Therefore, the WHO-5 should be considered as an outcome scale in Fava's well-being therapies.

In conclusion, the WHO-5 is a short questionnaire consisting of 5 simple and non-invasive questions, which tap into the subjective well-being of the respondents. The scale has adequate validity both as a screening tool for depression and as an outcome measure in clinical trials and has been applied successfully as a generic scale for well-being across a wide range of study fields. In our opinion, the findings of this review show that the WHO-5 is a highly useful tool that can be applied in both clinical practice (for instance to screen for depression) as well as in research studies in order to assess well-being over time or to compare well-being between groups.

Funding

This project was funded by the Psychiatric Research Unit, Psychiatric Centre North Zealand, Copenhagen University Hospital, Hillerød, Denmark.

Disclosure Statement

The authors declare no conflicts of interest.

References

- 1 Feinstein AR: Clinimetrics. New Haven. Yale University Press, 1987.
- 2 Tomba E, Bech P: Clinimetrics and clinical psychometrics: macro- and microanalysis. *Psychother Psychosom* 2012;81:333–343.
- 3 Fava GA: Clinical judgment in psychiatry. *Requiem or reveille?* *Nord J Psychiatry* 2013;67: 1–10.
- 4 Gill TM, Feinstein AR: A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;272:619–626.
- 5 Ware JE Jr: The status of health assessment 1994. *Annu Rev Public Health* 1995;16:327–354.
- 6 Bech P, Gudex C, Staehr Johansen K: The WHO (Ten) Well-Being Index: validation in diabetes. *Psychother Psychosom* 1996;65: 183–190.
- 7 Warr P, Banks M, Ullah P: The experience of unemployment among black and white urban teenagers. *Br J Psychol* 1985;76:75–87.
- 8 Staehr Johansen K: Multicentre Continuous Subcutaneous Infusion Pump Feasibility and Acceptability Study Experience. Copenhagen, WHO Regional Office for Europe, 1989.
- 9 Mokken RJ: *Theory and Practice of Scale Analysis*. Berlin, Mouton, 1971.
- 10 Bech P: *Rating Scales for Psychopathology, Health Status and Quality of Life. A Compendium on Documentation in Accordance with the DSM-III-R and WHO Systems*. Berlin, Springer, 1993.
- 11 Jahoda M: *Current Concepts of Positive Mental Health*. New York, Basic Books Publisher, 1958.
- 12 Bech P: Health-related quality of life measurements in the assessment of pain clinic results. *Acta Anaesthesiol Scand* 1999;43:893–896.
- 13 World Health Organization: *International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). Diagnostic Criteria for Research*. Geneva, World Health Organization, 1993.
- 14 Bech P, Wermuth L: Applicability and validity of the Major Depression Inventory in patients with Parkinson's disease. *Nord J Psychiatry* 1998;52:305–309.
- 15 Bech P: Quality of life and rating scales for depression; in Preskorn SH, Feighner JP, Stranga CY, Ross R (eds): *Antidepressants: Past, Present and Future*. New York, Springer, 2004, pp 148–170.

- 16 Staehr Johansen K: The use of well-being measures in primary health care – the DepCare project; in World Health Organization, Regional Office for Europe: Well-Being Measures in Primary Health Care – the DepCare Project. Geneva, World Health Organization, 1998, target 12, E60246.
- 17 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D: The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–e34.
- 18 Eurofound: European Quality of Life Survey 2012. 2014. <http://www.eurofound.europa.eu/surveys/eqls/2011/>.
- 19 Hall T, Krahn GL, Horner-Johnson W, Lamb G; Rehabilitation Research and Training Center Expert Panel on Health Measurement: Examining functional content in widely used Health-Related Quality of Life scales. *Rehabil Psychol* 2011;56:94–99.
- 20 Rasch, G: From Guttman's pioneer work to item response theory analysis (IRT); in Bech P (ed): *Clinical Psychometrics*. Oxford, Wiley-Blackwell, 2012.
- 21 Blom EH, Bech P, Hogberg G, Larsson JO, Gotzschue E: Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales – testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. *Health Qual Life Outcomes* 2012;10:149.
- 22 Lucas-Carrasco R, Allerup P, Bech P: The validity of the WHO-5 as an early screening for apathy in an elderly population. *Curr Gerontol Geriatr Res* 2012;2012:171857.
- 23 Bech P, Olsen L, Kjoller M, Rasmussen N: 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.
- 24 Ellervik C, Kvetny J, Christensen KS, Vestergaard M, Bech P: Prevalence of depression, quality of life and antidepressant treatment in the Danish General Suburban Population Study. *Nord J Psychiatry* 2014;68:507–512.
- 25 Birket-Smith M, Hansen BH, Hanash JA, Hansen JF, Rasmussen A: Mental disorders and general well-being in cardiology outpatients – 6-year survival. *J Psychosom Res* 2009;67:5–10.
- 26 Wade AG, Ford I, Crawford G, McMahon AD, Nir T, Laudon M, Zisapel N: Efficacy of prolonged release melatonin in insomnia patients aged 55–80 years: quality of sleep and next-day alertness outcomes. *Curr Med Res Opin* 2007;23:2597–2605.
- 27 Bech P, Lunde M, Bech-Andersen G, Lindberg L, Martiny K: Psychiatric outcome studies (POS): does treatment help the patients? A Popperian approach to research in clinical psychiatry. *Nord J Psychiatry* 2007;61(suppl 46):4–34.
- 28 Hoffman CJ, Ersser SJ, Hopkinson JB, Nicholls PG, Harrington JE, Thomas PW: Effectiveness of mindfulness-based stress reduction in mood, breast- and endocrine-related quality of life, and well-being in stage 0 to III breast cancer: a randomized, controlled trial. *J Clin Oncol* 2012;30:1335–1342.
- 29 Logtenberg SJ, Kleefstra N, Houweling ST, Groenier KH, Gans RO, Bilo HJ: Health-related quality of life, treatment satisfaction, and costs associated with intraperitoneal versus subcutaneous insulin administration in type 1 diabetes: a randomized controlled trial. *Diabetes Care* 2010;33:1169–1172.
- 30 Robinson SK, Viirre ES, Bailey KA, Gerke MA, Harris JP: Randomized placebo-controlled trial of a selective serotonin reuptake inhibitor in the treatment of nondepressed tinnitus subjects. *Psychosom Med* 2005;67:981–988.
- 31 Guico-Pabia CJ, Fayyad RS, Soares CN: Assessing the relationship between functional impairment/recovery and depression severity: a pooled analysis. *Int Clin Psychopharmacol* 2012;27:1–7.
- 32 Martiny K, Refsgaard E, Lund V, Lunde M, Sorensen L, Thougard B, Lindberg L, Bech P: A 9-week randomized trial comparing a chronotherapeutic intervention (wake and light therapy) to exercise in major depressive disorder patients treated with duloxetine. *J Clin Psychiatry* 2012;73:1234–1242.
- 33 Lowe B, Spitzer RL, Grafe K, Kroenke K, Quenter A, Zipfel S, Buchholz C, Witte S, Herzog W: Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord* 2004;78:131–140.
- 34 Nicolucci A, Rossi MC, Pellegrini F, Lucisano G, Pintaudi B, Gentile S, Marra G, Skovlund SE, Vespasiani G; BENCH-D Study Group: Benchmarking network for clinical and humanistic outcomes in diabetes (BENCH-D) study: protocol, tools, and population. Springerplus 2014; 3:83.
- 35 Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, Grosse Holtforth M: Measuring depression with a well-being index: further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. *J Affect Disord* 2014;156:240–244.
- 36 Andrews FM, Withey SB: *Social Indicators of Well-Being. Americans' Perceptions of Life Quality*. New York, Plenum Press, 1976.
- 37 Stefanello S, Cais CF, Mauro ML, Freitas GV, Botega NJ: Gender differences in suicide attempts: preliminary results of the Multisite Intervention Study on Suicidal Behavior (SU-PRE-MISS) from Campinas, Brazil. *Rev Bras Psiquiatr* 2008;30:139–143.
- 38 Sisask M, Varnik A, Kolves K, Konstabel K, Wasserman D: Subjective psychological well-being (WHO-5) in assessment of the severity of suicide attempt. *Nord J Psychiatry* 2008;62:431–435.
- 39 Vijayakumar L, Ali ZSS, Umamaheswari C: Socio cultural and clinical factors in repetition of suicide attempts: a study from India. *Int J Cult Ment Health* 2008;1:3–9.
- 40 Awata S, Bech P, Koizumi Y, Seki T, Kuriyama S, Hozawa A, Ohmori K, Nakaya N, Matsuoaka H, Tsuji I: Validity and utility of the Japanese version of the WHO-Five Well-Being Index in the context of detecting suicidal ideation in elderly community residents. *Int Psychogeriatr* 2007;19:77–88.
- 41 Hensing G, Holmgren K, Mardby AC: Harmful alcohol habits were no more common in a sample of newly sick-listed Swedish women and men compared with a random population sample. *Alcohol Alcohol* 2011;46:471–477.
- 42 Elholm B, Larsen K, Hornnes N, Zierau F, Becker U: Alcohol withdrawal syndrome: symptom-triggered versus fixed-schedule treatment in an outpatient setting. *Alcohol Alcohol* 2011;46:318–323.
- 43 Hoxmark E, Wynn TN, Wynn R: Loss of activities and its effect on the well-being of substance abusers. *Scand J Occup Ther* 2012;19:78–83.
- 44 Bergmann N, Ballegaard S, Holmager P, Kristiansen J, Gyntelberg F, Andersen LJ, Hjalmarson A, Bech P, Arendt-Nielsen L, Faber J: Pressure pain sensitivity: a new method of stress measurement in patients with ischemic heart disease. *Scand J Clin Lab Invest* 2013;73:373–379.
- 45 Garnefski N, Kraaij V, Schroevers MJ, Sommersen GA: Post-traumatic growth after a myocardial infarction: a matter of personality, psychological health, or cognitive coping? *J Clin Psychol Med Settings* 2008;15:270–277.
- 46 Schneider CB, Pilhatsch M, Rifati M, Jost WH, Wodarz F, Ebersbach G, Djundja D, Fuchs G, Gies A, Odin P, Reifschneider G, Wolz M, Bottesi A, Bauer M, Reichmann H, Storch A: Utility of the WHO-Five Well-being Index as a screening tool for depression in Parkinson's disease. *Mov Disord* 2010;25:777–783.
- 47 Jorgensen B, Friis GJ, Gottrup F: Pain and quality of life for patients with venous leg ulcers: proof of concept of the efficacy of Bia-tain-Ibu, a new pain reducing wound dressing. *Wound Repair Regen* 2006;14:233–239.
- 48 Volinn E, Yang B, He J, Sheng X, Ying J, Volinn W, Zhang J, Zuo Y: West China hospital set of measures in Chinese to evaluate back pain treatment. *Pain Med* 2010;11:637–647.
- 49 Verecke E, Susanszky E, Kopp M, Ratko I, Czibalmos A, Nagy Z, Palkonyai E, Hodinka L, Temesvari PI, Kiss E, Toro K, Poor G: Psychosocial, educational, and somatic factors in chronic nonspecific low back pain. *Rheumatol Int* 2013;33:587–592.
- 50 Cole R, Hayes B, Jones D, Shah S: Coping strategies used by school staff after a crisis: a research note. *J Loss Trauma* 2013;18:472–481.

- 51 Feicht T, Wittmann M, Jose G, Mock A, von Hirschhausen E, Esch T: Evaluation of a seven-week web-based happiness training to improve psychological well-being, reduce stress, and enhance mindfulness and flourishing: a randomized controlled occupational health study. *Evid Based Complement Alternat Med* 2013;2013:676953.
- 52 Gao J, Weaver SR, Dai J, Jia Y, Liu X, Jin K, Fu H: Workplace social capital and mental health among Chinese employees: a multi-level, cross-sectional study. *PLoS One* 2014; 9:e85005.
- 53 Nielsen K, Randall R: The importance of employee participation and perceptions of changes in procedures in a teamworking intervention. *Work Stress* 2012;26:91–111.
- 54 Schutte S, Chastang JF, Malard L, Parent-Thirion A, Vermeylen G, Niedhammer I: Psychosocial working conditions and psychological well-being among employees in 34 European countries. *Int Arch Occup Environ Health* 2014;87:897–907.
- 55 Jung J, Ernstmann N, Nitzsche A, Driller E, Kowalski C, Lehner B, Stieler-Lorenz B, Friepartner K, Schmidt A, Pfaff H: Exploring the association between social capital and depressive symptoms: results of a survey in German information and communication technology companies. *J Occup Environ Med* 2012;54: 23–30.
- 56 World Health Organization: *International Classification of Impairments, Disabilities, and Handicaps*. Geneva, World Health Organization, 1980.
- 57 Bech P: Measurement of psychological distress and well-being. *Psychother Psychosom* 1990;54:77–89.
- 58 Steffens DC, Helms MJ, Krishnan KR, Burke GL: Cerebrovascular disease and depression symptoms in the cardiovascular health study. *Stroke* 1999;30:2159–2166.
- 59 O'Reardon JP, Amsterdam JD: *Medical disorders and treatment-resistant depression*; in Amsterdam JD, Hornig M, Nierenberg AA (eds): *Treatment-Resistant Mood Disorder*. Cambridge, Cambridge University Press, 2001, pp 405–429.
- 60 Kan C, Silva N, Golden SH, Rajala U, Timonen M, Stahl D, Ismail K: A systematic review and meta-analysis of the association between depression and insulin resistance. *Diabetes Care* 2013;36:480–489.
- 61 Bergmann N, Gyntelberg F, Faber J: The appraisal of chronic stress and the development of the metabolic syndrome: a systematic review of prospective cohort studies. *Endocr Connect* 2014;3:R55–R80.
- 62 Ryff CD: Psychological well-being revisited: advances in the science and practice of eudaimonia. *Psychother Psychosom* 2014;83:10–28.
- 63 Fava GA: Well-being therapy: conceptual and technical issues. *Psychother Psychosom* 1999; 68:171–179.
- 64 Allgaier AK, Kramer D, Saravo B, Mergl R, Fejtikova S, Hegerl U: Beside the Geriatric Depression Scale: the WHO-Five Well-being Index as a valid screening tool for depression in nursing homes. *Int J Geriatr Psychiatry* 2013; 28:1197–1204.
- 65 Awata S, Bech P, Yoshida S, Hirai M, Suzuki S, Yamashita M, Ohara A, Hinokio Y, Matsuoka H, Oka Y: Reliability and validity of the Japanese version of the World Health Organization-Five Well-Being Index in the context of detecting depression in diabetic patients. *Psychiatry Clin Neurosci* 2007;61:112–119.
- 66 De Azevedo-Marques JM, Zuardi AW: COOP/WONCA charts as a screen for mental disorders in primary care. *Ann Fam Med* 2011;9:359–365.
- 67 Hammer EM, Hacker S, Hautzinger M, Meyer TD, Kubler A: Validity of the ALS-Depression-Inventory (ADI-12) – a new screening instrument for depressive disorders in patients with amyotrophic lateral sclerosis. *J Affect Disord* 2008;109:213–219.
- 68 Liwowsky I, Kramer D, Mergl R, Bramesfeld A, Allgaier A, Poppel E, Hegerl U: Screening for depression in the older long-term unemployed. *Soc Psychiatry Psychiatr Epidemiol* 2009;44:622–627.
- 69 Saipanish R, Lotrakul M, Sumrithe S: Reliability and validity of the Thai version of the WHO-Five Well-Being Index in primary care patients. *Psychiatry Clin Neurosci* 2009;63: 141–146.
- 70 Sibai AM, Chaaya M, Tohme RA, Mahfoud Z, AlAmin H: Validation of the Arabic version of the 5-item WHO Well Being Index in elderly population. *Int J Geriatr Psychiatry* 2009;24:106–107.
- 71 Allgaier AK, Pietsch K, Fruhe B, Prast E, Sigl-Glockner J, Schulte-Korne G: Depression in pediatric care: is the WHO-Five Well-Being Index a valid screening instrument for children and adolescents? *Gen Hosp Psychiatry* 2012;34:234–241.
- 72 Bonsignore M, Barkow K, Jessen F, Heun R: Validity of the five-item WHO Well-Being Index (WHO-5) in an elderly population. *Eur Arch Psychiatry Clin Neurosci* 2001;251:27–31.
- 73 Henkel V, Mergl R, Kohnen R, Maier W, Moller HJ, Hegerl U: Identifying depression in primary care: a comparison of different methods in a prospective cohort study. *BMJ* 2003;326:200–201.
- 74 Mergl R, Seidscheck I, Allgaier AK, Moller HJ, Hegerl U, Henkel V: Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. *Depress Anxiety* 2007; 24:185–195.
- 75 De Wit M, Pouwer F, Gemke RJB, Delemarre-Van De Waal HA, Snoek FJ: Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. *Diabetes Care* 2007;30: 2003–2006.
- 76 Furuya M, Hayashino Y, Tsujii S, Ishii H, Fukuhara S: Comparative validity of the WHO-5 Well-Being Index and two-question instrument for screening depressive symptoms in patients with type 2 diabetes. *Acta Diabetol* 2013;50:117–121.
- 77 Hajos TR, Pouwer F, Skovlund SE, Den Ouden BL, Geelhoed-Duijvestijn PH, Tack CJ, Snoek FJ: Psychometric and screening properties of the WHO-5 Well-Being Index in adult outpatients with type 1 or type 2 diabetes mellitus. *Diabet Med* 2013;30:e63–e69.
- 78 Barua A, Kar N: Screening for depression in elderly Indian population. *Indian J Psychiatry* 2010;52:150–153.
- 79 Bachner YG, Ayalon L: Initial examination of the psychometric properties of the short Hebrew version of the Zarit Burden Interview. *Aging Ment Health* 2010;14:725–730.