

# Validity of the SF-36 in Patients with Morbid Obesity

Tor-Ivar Karlsen<sup>a,b,c</sup> Einar K. Tveit<sup>a,c</sup> Gerd K. Natvig<sup>d</sup> Serena Tonstad<sup>e</sup> Jøran Hjeltnes<sup>a</sup>

<sup>a</sup> Morbid Obesity Center, Vestfold Hospital Trust, Tønsberg,

<sup>b</sup> Department of Psychosocial Health, University of Agder, Grimstad,

<sup>c</sup> Evjeklinikken AS, Evje,

<sup>d</sup> Department of Public Health and Primary Health Care, University of Bergen, Norway

<sup>e</sup> School of Public Health, Loma Linda University, Loma Linda, CA, USA

## Keywords

Lifestyle factors · Obesity · Psychological aspects ·  
Quality of life · SF-36 · Validity

## Summary

**Background:** The Medical Outcome Study 36-Item Short Form Health Survey (SF-36) is one of the most commonly applied generic quality of life instruments. The construct validity of the instrument in patients with morbid obesity is not established. **Participants and Methods:** A total of 475 morbidly obese patients (68% women) with a mean (standard deviation) age of 44.7 (11.8) years, weight of 123.5 (24.1) kg and BMI of 41.7 (6.3) kg/m<sup>2</sup>, who had been referred to a rehabilitation center, completed the SF-36 form. Exploratory factor analyses were performed to examine the underlying component structure of the questionnaire. Confirmatory factor analyses were performed to assess model fit. **Results:** The analysis suggested a 6-component structure rather than the 8-component structure used in the original SF-36. The first component consisted of items from the physical functioning subscale, the role physical subscale and the general health subscale, and explained 31% of the variance. The 6 components explained 61% of the total variance. The items loaded as expected in a physical and mental component. The assessment of model fit confirmed these findings. **Conclusion:** The 2 summary scales of the SF-36 have satisfactory validity in patients with morbid obesity. However, the validity of the 8 subscales is questionable, and the subscales should be interpreted with care.

## Introduction

Morbid obesity, defined as a BMI  $\geq 40$  kg/m<sup>2</sup> or BMI  $\geq 35$  kg/m<sup>2</sup> with comorbidities, is a chronic condition leading to premature morbidity and mortality [1]. A wide range of comorbidities accompany obesity, including impaired health-related quality of life (HRQL) [2]. HRQL is usually measured either with generic or disease-specific instruments. The generic instruments are designed to measure HRQL across divergent populations. The disease-specific instruments intend to measure HRQL aspects in a defined population. In clinical studies it is recommended to apply both types of instruments [3]. The Medical Outcome Study 36-Item Short Form Health Survey (SF-36) is a generic HRQL instrument developed in the USA and released for general use in the late 1980s [4]. The SF-36 consists of 36 questions (items) of which 35 are used to score the instrument. The SF-36 is one of the most popular generic instruments and has been used to assess HRQL in over 1,000 studies in the past 2 decades [5], including recent studies of obesity [6–14]. However, to our knowledge, only 1 study has examined the validity of the instrument in obese patients [15]. This study was performed as part of the QUOVADIS study, an observational multicenter study of 1,735 obese outpatients in Italy, of whom 40% had a BMI  $< 35$  kg/m<sup>2</sup>. A principal component analysis (PCA) with an oblique rotation suggested a 6-component solution instead of the standard 8-component solution, thereby questioning the construct validity of the 8 SF-36 subscales in patients with obesity. The study did not assess the structural validity of the 2 summary scales [15]. The aim of our study was to assess the structural validity of the SF-36 in a population of morbidly obese persons. We investigated whether the underlying component structure of the SF-36 supports a separation into 8 subscales and 2 summary scales in a sample of treatment-seeking morbidly obese patients.

**Table 1.** Descriptive anthropometric and socioeconomic characteristics of 475 morbidly obese treatment-seeking patients

Variable	Total	Women	Men	p <sup>a</sup>
Patients, n	475	326	149	
Age, years, mean (SD)	44.7 (11.8)	44.0 (11.6)	46.3 (12.2)	0.049
Weight, kg, mean (SD)	124 (24)	116 (20)	139 (25)	< 0.001
BMI, kg/m <sup>2</sup> mean (SD)	41.7 (6.3)	41.3 (6.1)	42.4 (6.8)	0.087
WHR, mean (SD)	0.98 (0.11)	0.92 (0.08)	1.10 (0.08)	< 0.001
Paid work, n (%)	283 (60)	198 (61)	85 (58)	0.612

<sup>a</sup>p value for gender differences.

SD = Standard deviation; WHR = waist/hip ratio.

## Participants and Methods

### Study Design, Setting, and Participants

This cross-sectional study was based on data collected at a Norwegian rehabilitation center specializing in treating morbidly obese patients. In order to reduce potential selection bias, all patients with a BMI  $\geq$  35 (n = 537) admitted to the rehabilitation center between May 2005 and September 2009 were sent the SF-36 form prior to their first admission. The study was conducted after written informed consent was obtained in accordance with the Helsinki protocol. The study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics (ref. 2010/954).

### Variables, Measurement, and Outcomes

The scale authors applied a PCA in the development of the subscales [16]. The 8 subscales are 'physical functioning' (10 items), 'role physical' (4 items), 'bodily pain' (2 items), 'general health' (5 items), 'vitality' (4 items), 'social functioning' (2 items), 'role emotional' (3 items), and 'mental health' (5 items). These 8 subscales then underwent a second order PCA which resulted in 2 summary scales; a 'physical summary scale' and a 'mental summary scale' [17].

### Study Size

There is little support in the literature in regard to sample size. Rules-of-thumb vary from 4 to 10 subjects per variable, with a minimum number of 200 subjects to ensure stability of the variance/covariance matrix [18]. De Vet et al. [19] concluded with 7 subjects per item, which with the 35 items in the SF-36 gives a minimum of 245 cases.

### Statistical Methods

Data are presented as mean (standard deviation, SD) or n (%) unless otherwise stated. Little's MCAR analyses [20] were performed prior to mean imputation of missing data. Possible differences between patients excluded from and included in the analyses were assessed using independent samples t-tests. To examine the component structure of the SF-36, we applied PCA with an oblique (promax) rotation [21, 22]. The Kaiser-Meyer-Olkin measure and Bartlett's test of sphericity were computed to determine whether the data in this sample were suitable for PCA. The decision of how many factors to retain is a critical component of exploratory factor analysis. Following Kaiser's criterion, eigenvalues of 1.0 were chosen to ensure that the extracted components accounted for a reasonably large proportion of the total variance [22]. Methodological research has shown that the Kaiser's criterion may include too many components [18, 23]. A parallel analysis [23] was conducted to affirm the extraction. Parallel analysis involves comparing the eigenvalues from the actual sample with those obtained from a randomly assigned dataset. A correlation matrix is computed from the randomly generated dataset, and the eigenvalues of the correlation matrix are computed. Components corresponding to the initial eigenvalues that are greater than the parallel average random eigenvalues should be retained. An item was considered to be loaded on a component if coefficients were 0.32 or larger [22]. To assess the model fit, we applied

confirmatory factor analyses (CFA) as structural equation modeling (SEM). 2 tests of goodness of fit were used to evaluate the models, the Comparative Fit Index (CFI) and the Root Mean Squared Error of Approximation (RMSEA). A CFI > 0.90 indicates that the model has a good fit to the data. An RMSEA < 0.08 indicates a good fit to the data, while values > 0.10 suggest that the model fit is unsatisfactory [22]. The statistical analyses were performed with SPSS Statistics v. 17.0 (SPSS Inc., Chicago, IL, USA) for data processing, PRELIS v. 8.8 (SSI Inc., Lincolnwood, IL, USA) for PCA, Windows MonteCarlo PCA for parallel analysis, and AMOS v. 18.0 (IBM Corp., Armonk, NY, USA) for CFA.

## Results

### Patients

All subjects (n = 537) were non-immigrants of Caucasian ethnicity. Of these, 62 (11%) patients had more than 60% missing item values [17], leaving 475 (89%) patients eligible for further statistical analysis. Table 1 displays various anthropometric and sociodemographic variables for the 475 included patients. Little's MCAR test showed that data was missing completely at random (p = 0.303). There were no statistically significant differences between gender, age, weight, BMI, and occupational status between the patients eligible for analyses and the patients excluded from analyses (all p > 0.05).

### Component Extraction

The data met the Kaiser-Meyer-Olkin measure (0.91) and the Bartlett's test of sphericity criterion (p < 0.001) for performing PCA. Following the Kaiser's criterion, 7 components had an eigenvalue > 1.0. When controlling the extraction in a parallel analysis, component 7 (eigenvalue 1.01) was rejected, and components 1–6 were retained.

### Exploratory Factor Analysis

The component loads are presented in table 2. Component 1 had an eigenvalue of 10.76 and an explained variance of 31%. The items belonging to the subscale 'physical functioning' split in their loads on components 1 and 2 (eigenvalue 4.61, variance 13.2%). The items of 'role physical' also loaded on component 1. The 3 'role emotional' items loaded on component 3 (eigenvalue 1.69, variance 4.8%). 'Bodily pain' items loaded on component 4 (eigenvalue 1.56, variance 4.4%). The items of the subscales 'social functioning' and 'mental health'

**Table 2.** Rotated structure in a 6-component PCA, oblique (promax) rotation; presenting item loads of the SF-36 on components 1–6 (items with loadings  $\geq 0.32$  in italics) for morbidly obese patients ( $n = 475$ )

Item	Item text	Subscale	Component					
			1	2	3	4	5	6
sf3a	'Health limits vigorous activities'	PF	<i>0.64</i>	-0.10	0.00	-0.07	-0.12	-0.04
sf3b	'Health limits moderate activities'		<i>0.58</i>	0.26	-0.10	0.05	0.07	-0.02
sf3c	'Health limits lifting and carrying groceries'		0.06	<i>0.49</i>	-0.02	0.22	0.06	-0.01
sf3d	'Health limits climbing several flights of stairs'		<i>0.76</i>	0.18	-0.02	-0.15	-0.06	-0.03
sf3e	'Health limits climbing one flight of stairs'		<i>0.40</i>	<i>0.55</i>	-0.05	-0.13	0.04	-0.01
sf3f	'Health limits bending, kneeling or stooping'		<i>0.76</i>	0.08	-0.05	-0.01	-0.09	-0.09
sf3g	'Health limits walking more than a mile'		<i>0.62</i>	<i>0.38</i>	0.05	-0.11	-0.12	-0.02
sf3h	'Health limits walking several hundred yards'		0.20	<i>0.80</i>	0.06	-0.05	-0.03	0.05
sf3i	'Health limits walking one hundred yards'		-0.10	<i>0.90</i>	0.07	0.02	0.00	0.05
sf3j	'Health limits bathing or dressing yourself'		0.18	<i>0.39</i>	-0.14	0.05	0.10	0.05
sf4a	'Due to physical health cut down time on work'	RP	0.25	0.14	0.13	0.26	0.04	0.06
sf4b	'Due to physical health accomplished less'		<i>0.45</i>	0.00	0.21	0.13	-0.13	0.18
sf4c	'Due to physical health limited in work'		<i>0.53</i>	0.00	-0.03	0.27	0.01	-0.08
sf4d	'Due to physical health difficulty performing work'		<i>0.61</i>	0.02	0.03	0.15	0.01	-0.08
sf7	'Bodily pain'	BP	0.07	0.01	-0.01	<i>0.79</i>	-0.09	0.08
sf8	'Pain interfered with normal work'		0.13	-0.01	0.01	<i>0.85</i>	-0.01	0.00
sf1	'Self reported health'	GH	<i>0.47</i>	0.13	-0.04	0.09	0.17	-0.04
sf11a	'Seem to get ill more easily than other people'		0.18	-0.09	0.04	0.14	0.26	-0.08
sf11b	'As healthy as anyone I know'		0.28	-0.04	-0.02	0.16	0.21	-0.05
sf11c	'Expect my health to get worse'		0.21	-0.05	0.00	-0.06	0.04	0.15
sf11d	'Health is excellent'		<i>0.47</i>	-0.12	-0.09	0.06	0.15	0.08
sf9a	'Felt full of life'	VT	<i>0.32</i>	-0.09	0.09	-0.09	0.07	0.26
sf9e	'Have had a lot of energy'		0.28	-0.06	0.07	-0.01	0.12	<i>0.40</i>
sf9g	'Felt worn out'		-0.07	0.05	-0.09	0.10	0.15	<i>0.79</i>
sf9i	'Felt tired'		-0.11	0.06	0.00	-0.01	0.06	<i>0.89</i>
sf6	'Extent physical health or emotional problems interfered with normal social activities'	SF	0.02	0.04	0.03	0.03	<i>0.80</i>	-0.03
sf10	'Physical health or emotional problems have interfered with social activities'		0.07	0.06	0.06	0.02	<i>0.76</i>	-0.06
sf5a	'Due to emotional problems cut time on work'	RE	-0.01	0.05	<i>0.61</i>	0.01	<i>0.33</i>	-0.08
sf5b	'Due to emotional problems accomplished less'		-0.04	0.03	<i>0.82</i>	0.00	0.14	0.02
sf5c	'Due to emotional problems work less carefully'		0.01	0.01	<i>0.63</i>	0.00	0.20	-0.06
sf9b	'Been very nervous'	MH	-0.16	0.09	0.07	-0.05	<i>0.69</i>	0.09
sf9c	'Felt so down in the dumps'		-0.07	-0.02	-0.07	0.01	<i>0.91</i>	-0.03
sf9d	'Felt calm and peaceful'		0.03	-0.02	0.12	-0.09	<i>0.67</i>	0.03
sf9f	'Felt downhearted and low'		-0.10	-0.05	-0.09	-0.04	<i>0.86</i>	0.13
sf9h	'Been happy'		0.15	-0.14	0.09	-0.06	<i>0.53</i>	0.08

PF = Physical functioning; RP = role physical; BP = bodily pain; GH = general health; VT = vitality; SF = social functioning; RE = role emotional; MH = mental health.

loaded on component 5 (eigenvalue 1.40, variance 4%), and 'vitality' on component 6 (eigenvalue 1.26, variance 3.6%). The 6 components explained 61% of the total variance. The 2-component analysis (table 3) confirmed the grouping of the physically oriented items in one component and the mentally oriented items into the other. The exception was a split of the 'general health' subscale where 3 of the 5 items loaded – 2 on the physical component and 1 on the mental component.

#### Confirmatory Factor Analysis

Four models were tested: the 6-component PCA model, the original 8-component model, the single item PCA 2-component model, and the original second order 2-component model. The structural equation modeling (SEM) analyses confirmed the PCA. The 6-component PCA model had better fit (RMSEA = 0.08, CFI = 0.81) than the original 8-component model (RMSEA = 0.13, CFI = 0.47). The analysis of the

**Table 3.** Rotated structure in a 2-component PCA, oblique (promax) rotation; presenting item loads of the SF-36 on components 1 and 2 (items with loadings  $\geq 0.32$  in italics) for morbidly obese patients (n = 475)

Item	Item text	Composite score	Component	
			1	2
sf3a	'Health limits vigorous activities'	physical dimension	<i>0.40</i>	-0.01
sf3b	'Health limits moderate activities'		<i>0.76</i>	0.00
sf3c	'Health limits lifting and carrying groceries'		<i>0.62</i>	-0.07
sf3d	'Health limits climbing several flights of stairs'		<i>0.70</i>	-0.05
sf3e	'Health limits climbing one flight of stairs'		<i>0.75</i>	-0.14
sf3f	'Health limits bending, kneeling or stooping'		<i>0.70</i>	-0.07
sf3g	'Health limits walking more than a mile'		<i>0.78</i>	-0.14
sf3h	'Health limits walking several hundred yards'		<i>0.82</i>	-0.17
sf3i	'Health limits walking one hundred yards'		<i>0.71</i>	-0.21
sf3j	'Health limits bathing or dressing yourself'		<i>0.55</i>	-0.05
sf4a	'Due to physical health cut down time on work'		<i>0.47</i>	0.23
sf4b	'Due to physical health accomplished less'		<i>0.45</i>	0.24
sf4c	'Due to physical health limited in work'		<i>0.60</i>	0.08
sf4d	'Due to physical health difficulty performing work'		<i>0.61</i>	0.11
sf7	'Bodily pain'	<i>0.51</i>	0.13	
sf8	'Pain interfered with normal work'	<i>0.57</i>	0.21	
sf1	'Self reported health'	<i>0.55</i>	0.18	
sf11a	'Seem to get ill more easily than other people'	<i>0.14</i>	<i>0.32</i>	
sf11b	'As healthy as anyone I know'	<i>0.28</i>	<i>0.26</i>	
sf11c	'Expect my health to get worse'	<i>0.09</i>	<i>0.18</i>	
sf11d	'Health is excellent'	<i>0.32</i>	<i>0.28</i>	
sf9a	'Felt full of life'	<i>0.11</i>	<i>0.38</i>	
sf9e	'Have had a lot of energy'	<i>0.16</i>	<i>0.51</i>	
sf9g	'Felt worn out'	<i>0.09</i>	<i>0.57</i>	
sf9i	'Felt tired'	<i>0.01</i>	<i>0.57</i>	
sf6	'Extent physical health or emotional problems interfered with normal social activities'	mental dimension	0.03	<i>0.79</i>
sf10	'Physical health or emotional problems have interfered with social activities'		0.08	<i>0.76</i>
sf5a	'Due to emotional problems cut time on work'		-0.01	<i>0.67</i>
sf5b	'Due to emotional problems accomplished less'		-0.06	<i>0.68</i>
sf5c	'Due to emotional problems work less carefully'		-0.04	<i>0.59</i>
sf9b	'Been very nervous'		-0.11	<i>0.72</i>
sf9c	'Felt so down in the dumps'		-0.10	<i>0.82</i>
sf9d	'Felt calm and peaceful'		-0.10	<i>0.78</i>
sf9f	'Felt downhearted and low'		-0.17	<i>0.86</i>
sf9h	'Been happy'		-0.09	<i>0.73</i>

PF = Physical functioning, RP = role physical, BP = bodily pain, GH = general health, VT = vitality, RE = role emotional, SF = social functioning, MH = mental health.

single item 2-component PCA model showed an RMSEA of 0.08 and a CFI of 0.84. The analysis of the second order original model showed an RMSEA of 0.07 and a CFI of 0.88.

## Discussion

### Key Results

The items of the SF-36 loaded on 6 main components when tested in a sample of morbidly obese patients. When forcing the PCA into a 2-component extraction, the items loaded into a physically and a mentally oriented domain. The analyses of model fit show an acceptable fit for the 6-component model

and an unacceptable fit for the original 8-component model. The 2-component model has acceptable fit, both as a second order model and as a single item model. These findings have 2 implications. Firstly, the original subscales of the SF-36 should be interpreted with care in this patient group. Secondly, the structural validity of the 2 summary scales is satisfactory, indicating that the summary scales can be preferred to the subscales when assessing HRQL in morbidly obese patients initiating lifestyle treatment.

### Components 1 and 2, Physical Functioning, and Role Physical

The items belonging to the 'physical functioning' subscale in normal subjects loaded on 2 components in our sample,

thereby diminishing the structural validity of this particular subscale. This is in correspondence with earlier findings [15]. In our data, component 1 consisted of 5 out of 10 ‘physical functioning’ items together with items from the ‘role physical subscale’. The items loading on component 1 are focusing on limitations in moderate and vigorous activities together with role limitations. The items loading on component 2 are focusing on severe physical problems like problems carrying groceries, climbing 1 flight of stairs and bathing or dressing. There is robust evidence of a relationship between obesity and impaired physical health [24–28]. A possible interpretation of this grouping of items could be that most physically impaired obese patients interpret the items differently than the obese patients with less physical impairment [15]. If this interpretation is correct, the ‘physical functioning’ subscale should be used with care in studies of morbidly obese patients.

#### *Component 3, Role Emotional*

In accordance with the scale authors, the ‘role emotional’ subscale intends to measure problems with work or other daily activities as a result of emotional problems. The analysis supports the validity of this subscale.

#### *Component 4, Bodily Pain*

In our data of obese subjects, the 2 items forming the ‘bodily pain’ subscale loaded together on component 4. Osteoarthritis, knee pain, and back pain are common among obese individuals and are positively associated with BMI [24–28]. Higher BMI is shown to be strongly associated with impaired scores on the ‘bodily pain’ subscales [15]. The uniform loading of this items strengthen the validity of this subscale in morbid obesity.

#### *Component 5, Mental Health, and Social Functioning*

Items belonging to the 2 subscales ‘mental health’ and ‘social functioning’ grouped together in 1 component in our study. This component may be interpreted as a ‘psychosocial domain’ that may be relevant for the study of the morbidly obese. In a literature review of obesity and depression, Markowitz et al. [29] found evidence of a relationship between these 2 conditions. The authors propose that behavioral mechanisms such as functional impairment and repeated dieting, cognitive mechanisms such as body image dissatisfaction and poor self-rated health, and social mechanisms such as stigma may all play a role in the pathway between obesity and depression [29]. Likewise, Fabricatore and Wadden [30] indicated a relationship between severe obesity, psychiatric conditions, and mood disturbances, and that this may have both social implications and social causes. These relationships [29, 30] may explain the grouping of items into a ‘psychosocial domain’ in this sample of morbidly obese patients. In morbid obesity, these 2 subscales should be interpreted together.

#### *Component 6, Vitality*

Three out of 4 items belonging to the ‘vitality’ subscale in normal subjects loaded together on the same component in our obese patients. The ‘vitality’ subscale intends to measure energy level and fatigue. Morbid obesity is a condition associated with chronic fatigue [31], and this grouping of items may indicate that the ‘vitality’ subscale is a relevant single HRQL domain in morbidly obese patients.

#### *Summary Scales*

Our study is the first to assess the validity of the 2 summary scales in morbidly obese patients. The scale authors applied a second order PCA to extract the 2 summary scales from the 8 subscales [17]. With regard to the weakened structural validity of the original subscales in this population, we have taken the single items into account in an oblique (promax) rotated PCA. This model was tested against the original second order component model. The SEM analyses showed a moderate but acceptable fit of both 2-component models – the original second order model and the single item 2-component model. The main pattern is the same in our analyses compared to the scale authors’ analysis. The mentally oriented items loaded on one component and physically oriented items loaded on the other, implying acceptable structural validity of the summary scales in morbid obesity.

#### *Concluding Remarks*

Our results do not dissuade the use of SF-36 in studies of morbid obesity, but there are some important considerations and limitations to the use of SF-36 as an outcome measure in this population. Firstly, the clinical interpretation of the subscale scores is obscured due to ambiguity regarding their meaning. Researchers utilizing the SF-36 subscales should be aware of this and use, as recommended [3], obesity-specific HRQL instruments as supplementary measures. The subscale scores of the SF-36 will probably be more adequately interpreted when compared to such diagnosis-specific instruments. Secondly, despite the high validity of the summary scales, there are disadvantages to using them at the expense of the subscales. The most important is related to the loss of information. The ‘physical summary scale’ consists of 21 items and the ‘mental summary scale’ of 14 items. Consequently, a lot of detailed information is lost. Another disadvantage of using the summary scales is related to the scoring algorithm [17]. Some authors claim that the scoring algorithm produce unreliable results [32, 33]. The scale authors argue as a response that the scoring algorithm is for all practical purposes reasonable [34]. Again, other authors propose to adjust the scoring algorithm [35]. As a consequence of this discourse, researchers planning to use the summary scales in the study of obesity should examine the literature on which algorithm to use and comment on their choice.

The strength of this study is that it is based on a large sample consisting of consecutive morbidly obese patients referred

from public general hospitals, and the patients are as such expected to be representative of lifestyle treatment-seeking morbidly obese patients. However, the study has some limitations. Studies conducted in the early 1990s indicate that severely obese persons seeking treatment have significantly greater psychological stress than the non-treatment seeking severely obese [30]. Generalizability to the morbidly obese population may consequently be limited. Nevertheless, as the incidence of morbid obesity has increased over the last 2 decades, it is to be expected that larger groups of morbidly obese today wish to receive treatment, diminishing the potential discrepancy between sample results and population. To further increase generalizability, a confirmatory factor analysis should be performed on a sample independent of the one used in the exploratory factor analysis. We have used the same sample in both analyses, and we consider this as a limitation of this study. Another weakness of this study is that the patients are all Caucasians and that all patients were admitted for lifestyle treatment. As such, the conclusions may not be valid for patients seeking surgery or patients from other ethnic groups.

## Conclusions

The summary scales of the SF-36 have satisfactory construct validity in patients with morbid obesity but the subscale scores should be interpreted with care. The summary scales of the SF-36 may be considered as the primary scales when assessing HRQL in morbidly obese patients initiating lifestyle treatment.

## Authors' Contributions

TIK conceived the study design, collected data, conducted the statistical analyses and drafted the manuscript. EKT assessed the statistical outcomes. GKN, ST, and JH revised and approved the final manuscript.

## Disclosure Statement

The authors declare that they have no competing interests.

## References

- World Health Organisation: The World Health Report 2002 – Reducing Risks, Promoting Healthy Life. Geneva, WHO, 2002.
- Bray GA: Medical consequences of obesity. *J Clin Endocrinol Metab* 2004;89:2583–9.
- Fontaine KR, Barofsky I: Obesity and health-related quality of life. *Obes Rev* 2001;2:173–82.
- Ware JE, Jr., Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
- Ware JE Jr: SF-36 health survey update. *Spine (Phila Pa 1976)* 2000;25:3130–9.
- Ackermann RT, Edelstein SL, Narayan KM, et al.: Changes in health state utilities with changes in body mass in the Diabetes Prevention Program. *Obesity (Silver Spring)* 2009;17:2176–81.
- Adams TD, Pendleton RC, Strong MB, et al.: Health outcomes of gastric bypass patients compared to nonsurgical, nonintervened severely obese. *Obesity (Silver Spring)* 2010;18:121–30.
- Anandacoomarasamy A, Caterson ID, Leibman S, et al.: Influence of BMI on health-related quality of life: comparison between an obese adult cohort and age-matched population norms. *Obesity (Silver Spring)* 2009;17:2114–8.
- Burton PR, Brown W, Laurie C, et al.: Outcomes, satiety, and adverse upper gastrointestinal symptoms following laparoscopic adjustable gastric banding. *Obes Surg* 2011;21:574–81.
- Corsonello A, Lucchetti M, Corica F, Apolone G, Marchesini G: Metabolic syndrome and health-related quality of life: does psychological well-being matter? *Ann Epidemiol* 2008;18:592–3.
- Huang CJ, Hu HT, Fan YC, Liao YM, Tsai PS: Associations of breakfast skipping with obesity and health-related quality of life: evidence from a national survey in Taiwan. *Int J Obes (Lond)* 2010;34:720–5.
- Mond JM, Baune BT: Overweight, medical comorbidity and health-related quality of life in a community sample of women and men. *Obesity (Silver Spring)* 2009;17:1627–34.
- Riesco E, Rossel N, Rusques C, et al.: Impact of weight reduction on eating behaviors and quality of life: influence of the obesity degree. *Obes Facts* 2009;2:87–95.
- Wee HL, Wu Y, Thumboo J, Lee J, Tai ES: Association of body mass index with Short-Form 36 physical and mental component summary scores in a multiethnic Asian population. *Int J Obes (Lond)* 2010;34:1034–43.
- Corica F, Corsonello A, Apolone G, Lucchetti M, Melchionda N, Marchesini G: Construct validity of the Short Form-36 Health Survey and its relationship with BMI in obese outpatients. *Obesity (Silver Spring)* 2006;14:1429–37.
- Ware JE Jr, Kosinski M, Gandek B: SF-36 Health Survey: Manual & Interpretation Guide. Lincoln, RI, QualityMetric Incorporated, 1993, 2000, 2005.
- Ware JE, Kosinski M: SF-36 Physical and Mental Health Summary Scales: A Manual for Users of Version 1, 2nd ed. Lincoln, RI, QualityMetric Incorporated, 2001.
- Kline P: *The Handbook of Psychological Testing*, 2nd ed. New York, NY, Routledge, 2000.
- De Vet HC, Ader HJ, Terwee CB, Pouwer F: Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Qual Life Res* 2005;14:1203–18.
- Hill M: SPSS Missing Value Analysis™ 7.5. Chicago, IL, SPSS Inc., 1997.
- Reed PJ, Moore DD: SF-36 as a predictor of health states. *Value Health* 2000;3:202–7.
- Tabachnick B, Fidell LS: *Using Multivariate Statistics*, 4th ed. Boston, MA, Allyn & Bacon, 2006.
- Hayton JC, Allen DG, Scarpello V: Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Org Res Methods* 2004;7:191–205.
- Lementowski PW, Zelicoff SB: Obesity and osteoarthritis. *Am J Orthop (Belle Mead NJ)* 2008;37:148–51.
- Larsson UE, Mattsson E: Perceived disability and observed functional limitations in obese women. *Int J Obes Relat Metab Disord* 2001;25:1705–12.
- Kral JG: Morbidity of severe obesity. *Surg Clin North Am* 2001;81:1039–61.
- Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK: Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. *BMC Musculoskelet Disord* 2008;9:132.
- Barofsky I, Fontaine KR, Cheskin LJ: Pain in the obese: impact on health-related quality-of-life. *Ann Behav Med* 1997;19:408–10.
- Markowitz S, Friedman MA, Arent SM: Understanding the relation between obesity and depression: causal mechanisms and implications for treatment. *Clin Psychol* 2008;15:1–20.
- Fabricatore AN, Wadden TA: Psychological functioning of obese individuals. *Diabetes Spectrum* 2003;16:245–52.
- Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A: The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med* 1994;121:953–9.
- Taft C, Karlsson J, Sullivan M: Editorial comment: reply to Drs Ware and Kosinski. *Qual Life Res* 2001;10:415–20.
- Taft C, Karlsson J, Sullivan M: Do SF-36 summary component scores accurately summarize subscale scores? *Qual Life Res* 2001;10:395–404.
- Ware JE, Kosinski M: Interpreting SF-36 summary health measures: a response. *Qual Life Res* 2001;10:405–13.
- Farivar S, Cunningham W, Hays R: Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.1. *Health Qual Life Outcomes* 2007;5:54.