

RESEARCH

Open Access



Correlated physical and mental health composite scores for the RAND-36 and RAND-12 health surveys: can we keep them simple?

John Roger Andersen^{1,2,7*} , Kyrre Breivik^{1,3}, Inger Elise Engelund¹, Marjolein M. Iversen^{1,2}, Jorunn Kirkeleit^{1,4}, Tone Merete Norekvål^{1,2,5,6}, Kjersti Oterhals^{1,6} and Anette Storesund¹

Abstract

Background: The RAND-36 and RAND-12 (equivalent to versions 1 of the SF-36 Health Survey and SF-12 Health Survey, respectively) are widely used measures of health-related quality of life. However, there are diverging views regarding how to create the physical health and mental health composite scores of these questionnaires. We present a simple approach using an unweighted linear combination of subscale scores for constructing composite scores for physical and mental health that assumes these scores should be free to correlate. The aim of this study was to investigate the criterion validity and convergent validity of these scores.

Methods: We investigated oblique and unweighted RAND-36/12 composite scores from a random sample of the general Norwegian population ($N = 2107$). Criterion validity was tested by examining the correlation between unweighted composite scores and weighted scores derived from oblique principal component analysis. Convergent validity was examined by analysing the associations between the different composite scores, age, gender, body mass index, physical activity, rheumatic disease, and depression.

Results: The correlations between the composite scores derived by the two methods were substantial ($r = 0.97$ to 0.99) for both the RAND-36 and RAND-12. The effect sizes of the associations between the oblique versus the unweighted composite scores and other variables had comparable magnitudes.

Conclusion: The unweighted RAND-36 and RAND-12 composite scores demonstrated satisfactory criterion validity and convergent validity. This suggests that if the physical and mental composite scores are free to be correlated, the calculation of these composite scores can be kept simple.

Keywords: RAND-36, RAND-12, SF-36, SF-12, PCS, MCS, Oblique, Unweighted, Psychometric properties, Validity

Background

The RAND-36, and its brief version, the RAND-12 (equivalent to version 1 of the SF-36 Health Survey and SF-12 Health Survey, respectively), are freely available and widely used measures of generic health-related

quality of life (HRQoL) [1–4]. HRQoL refers to “how health impacts on an individual’s ability to function and his or her perceived well-being in physical, mental and social domains of life” [3]. The RAND-36/12 provides data on eight subscale scores and two composite scores of physical and mental health. The use of composite scores has become quite popular as they can simplify the interpretation of the findings [2]. However, despite the widespread use of the RAND-36/12 composite scores,

*Correspondence: johnra@hvl.no

¹ Department of Research and Development, Centre on Patient-Reported Outcomes, Bergen Hospital Trust, Bergen, Norway
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the choice of method for constructing them has been a controversial issue for decades [5–10].

Originally, Ware et al. [2, 11] provided algorithms for constructing composite scores based on orthogonal principal component analysis (PCA) to create a physical composite summary (PCS) and a mental composite summary (MCS) for version 1 of the SF-36/12 Health Surveys. They aimed to create pure PCS and MCS scores with little overlapping variance. To achieve this, all the scales/items must be included in the two composite scores but have different weights. However, this orthogonal approach has been criticized for producing inconsistencies between the composite scores and the observed data [5–7, 9].

Thus, a range of alternative scoring algorithms has been developed that do not restrict the correlations between the PCS and MCS [12]. One of the best documented alternatives to the orthogonal PCS and MCS was published by Farivar et al. in 2007, using oblique PCA to create the RAND-36/12 composite scores, which allowed correlations between them [7]. Overall, approaches such as this seem to be less prone to produce inconsistencies with the observed data [5–7].

On the other hand, a correlated PCS and MCS might not be without limitations. For example, a PCS and MCS from oblique PCA tend to be very strongly correlated, inducing multicollinearity [8]. Another issue is that the weights from oblique PCA fluctuate according to sample characteristics, making standardization across samples problematic [12, 13]. Furthermore, several authors have advocated the use of weights from confirmatory factor analysis (CFA) to create a PCS and MCS that are permitted to correlate [14, 15]. However, using CFA to construct composite scores can be problematic from a theoretical point of view, as a composite score, by nature, is a multi-dimensional construct [13, 16].

Hence, there are many alternatives to an orthogonal PCS and MCS for the RAND-36/12, making it unclear for researchers to decide which one to use [6, 12]. It has been argued that we often tend to make HRQoL scores unnecessarily complicated, for example, by using weighted scores [17]. Thus, simple unweighted composite scores have been proposed for the RAND-36/12. Such simple composite scores show promising criterion validity [5, 18]. This is not surprising given the strong correlations among the indicators of the RAND-36/12 composite scores. Weighting is probably of little value under such conditions [19], a logic that also applied when the original developers decided not to use weights for the SF-36 subscores [20]. However, data on the convergent validity for the unweighted RAND-36/12 is lacking in studies that have used unweighted composite scores for them [5, 18]. This is a limitation, as convergent validity is a crucial part of evaluating psychometric properties [21].

We present a simple approach to construct unweighted composite scores for the RAND-36/12, which implicitly assumes that these scores should be allowed to correlate. The composite scores were created by a linear combination of (1) The four subscales that have shown to be primarily indicative of physical health (physical functioning, physical role functioning, bodily pain and, general health), and (2) The four subscales that have shown to be primarily indicative of mental health (vitality, social functioning, emotional role functioning, and mental health). The aim of this study was to investigate the criterion and convergent validity of these scores by comparing them to established oblique composite scores. We hypothesized that unweighted and oblique composite scores would be highly correlated with each other and demonstrate equal convergent validity.

Methods

Design and study participants

We reused data from a representative survey of the general population of Norwegian adults aged 18–79 years. The methods have been described in detail previously [22]. In brief, the sample consisted of 2107 persons (36% response rate) who completed the Norwegian version of the RAND-36 (equivalent to the SF-36 version 1) as a postal questionnaire in 2015. All the items in the RAND-12 were taken directly from the RAND-36.

Demographic and other variables

We included self-reported data on age (10-year intervals), gender (women, men), marital/cohabitation status (no, yes), education (elementary school, high school, university < 4 years, and university \geq 4 years), strenuous physical activity habits (never, less than 1 h per week, 1–2 h per week, and \geq 3 h per week), self-reported height and weight (body mass index), and self-reported history of being diagnosed with a rheumatic disease or depression (no, yes) [22].

RAND-36 measures and scoring

Oblique RAND-36 PCS and MCS composite scores were created using the method, including the scoring coefficients described by Farivar et al. [7]. First, all the items were standardized into 0–100 scores. Second, eight subscales were created based on the mean scores of items belonging to the same scale: physical functioning (10 items), physical role functioning (4 items), bodily pain (2 items), general health (5 items), vitality (4 items), social functioning (2 items), emotional role functioning (3 items), and mental health (5 items). The subscale scores ranged from 0 to 100, with higher scores indicating better HRQoL. Third, eight z-scores were made using the mean and standard deviations of SF-36 subscales from

a 1998 US norm population, described in the manual of Ware et al. [2]. Forth, the z-scores were weighted using published scoring coefficients from oblique factor analysis. The only difference from the method of Ware et al. [2] and the present one is that we weighted the composite scores based on published scoring coefficients derived from the oblique PCA analysis in the study of Farivar et al. [7]. Finally, T-scores were made with a mean of 50 (SD = 10) representing the average scores in the US norm population [2].

The unweighted RAND-36 PCS and MCS composite scores were based on the original subscales, ranging from 0 to 100. Previous studies have shown that four subscale scores predominantly reflect physical health, while four others predominantly reflect mental health [6]. Thus, the unweighted RAND-36 PCS was created by adding the subscale scores for physical functioning, physical role functioning, bodily pain, and general health and dividing the sum by 4. The unweighted RAND-36 MCS was created by adding the subscale scores for vitality, social functioning, emotional role functioning, and mental health and dividing the sum by 4. This is quite similar to the RAND-HSI scoring but without the weights [6]. The unweighted RAND-36 PCS and MCS ranged from 0 to 100, with higher scores indicating better HRQoL.

RAND-12 measures and scoring

The oblique RAND-12 PCS and MCS composite scores were also created by the method of Farivar et al. [7]. The scoring was based on regressing the oblique RAND-36 PCS and MCS T-scores in separate models for the RAND-12 items. From these results, weighted dummy variables were used to create RAND-12 PCS and MCS T-scores, with higher scores indicating better HRQoL.

The unweighted RAND-12 PCS and MCS composite scores were created by standardizing the 12-items to 0–100 scores, in the same way as done for the RAND-36. Eight subscales were created, based on mean scores of items belonging to the same scale: physical functioning (2 items), physical role functioning (2 items), bodily pain (1 item), general health (1 item), vitality (1 item), social functioning (1 item), emotional role functioning (2 items), and mental health (2 items). Subscale scores ranged from 0 to 100, with higher scores indicating better HRQoL. The unweighted RAND-12 PCS score was created by adding the subscale scores for physical functioning, physical role functioning, bodily pain, and general health and dividing the sum by 4. The unweighted RAND-12 MCS score was created by adding the subscale scores for vitality, social functioning, emotional role functioning, and mental health and dividing the sum by 4. The unweighted RAND-12 PCS and MCS scores

ranged from 0 to 100, with higher scores indicating better HRQoL.

Statistics

Characteristics of the study population are presented as means and standard deviations or raw numbers and percentages. Descriptive statistics of the RAND-36/12 composite scores were based on z-scores (means = 0 and standard deviations = 1) because the different versions of the composite scores are not on the same metric. We present features of score distributions by medians, modes, floor, and ceiling effects (percentages), item-total correlations corrected for overlap, along with values for skewness and kurtosis. Floor or ceiling effects might be a problem if $\geq 15\%$ of respondents obtain the worst or best possible score [23]. The corrected item-total correlations were calculated for the unweighted RAND-36/12 composite scores: PCS (physical functioning, physical role functioning, bodily pain, and general health) and MCS (vitality, social functioning, emotional role functioning, and mental health), with values ≥ 0.4 indicating scores that consist of highly correlated variables [20]. Values of skewness ≥ 2 and kurtosis ≥ 7 suggest distributions of scores that begin to depart substantially from normality [24]. Associations between subscale scores and the PCS and MCS composite scores were examined using Pearson correlations or Spearman rank correlations. Criterion validity was examined by Pearson correlations between unweighted composite scores and scores derived from the oblique factor scoring coefficients [7]. Based on previous research and theory, the correlations between the unweighted and oblique RAND-36/12 composite scores measuring the same construct should be ≥ 0.95 [5, 18, 21]. Convergent validity was examined using Spearman rank coefficients between the composite scores and variables known to be related to HRQoL: age (years, continuous); sex (women = 0, men = 1); body mass index (units, continuous); physical activity (strenuous physical activity: never = 0, less than 1 h per week = 1, 1–2 h per week = 2, ≥ 3 h per week = 3), rheumatic disease (no = 0, yes = 1), and depression (no = 0, yes = 1) [2, 11, 25–28]. A correlation ≥ 0.2 regarding convergent validity suggest an effect size that might be of practical importance [29]. SPSS version 27 was used to perform the statistical analyses (IBM Corporation).

Results

The characteristics of the study participants are presented in Table 1. Table 2 shows that the means and standard deviations of unweighted RAND-36 versus RAND-12 scores (0–100) were not directly comparable, with the RAND-12 composite scores being somewhat lower. Descriptive statistics of the RAND-36/12 composite

Table 1 Characteristics of the study population ($n = 2107$)

Variables	Values
Age in years, n (%)	
18–29	105 (5.0)
30–39	203 (9.6)
40–49	400 (19.0)
50–59	484 (23.0)
60–69	519 (24.6)
70–79	396 (18.8)
Gender, n (%)	
Women	1143 (54.8)
Men	943 (45.2)
Married or cohabiting, n (%)	
Yes	1603 (76.1)
No	504 (23.9)
Education, n (%)	
Elementary school	79 (18.0)
High School	777 (37.0)
University < 4 years	457 (21.8)
University \geq 4 years	486 (23.2)
Strenuous physical activity, n (%)	
Never	292 (17.4)
Less than 1 h per week	375 (22.4)
1–2 h per week	596 (35.6)
\geq 3 h per week	410 (24.5)
Body mass index, mean (SD)	26.1 (4.5)
History of rheumatic disease, n (%)	
No	1766 (92.3)
Yes	148 (7.7)
History of depression, n (%)	
No	1656 (86.1)
Yes	268 (13.9)

Variables with fewer than 2017 observations are due to missing data

Table 2 Unweighted RAND 36/12 composite scores and subscores (0–100) presented with means and standard deviations

Scores	RAND-36	RAND-12
<i>Physical Composite Summary</i>	77.4 (21.5)	75.1 (23.4)
Physical functioning	87.3 (18.3)	86.9 (23.9)
Physical role functioning	76.8 (37.2)	74.4 (39.9)
Bodily pain	73.7 (25.1)	81.6 (25.1)
General health	71.8 (21.3)	57.3 (23.9)
<i>Mental Composite Summary</i>	79.4 (16.6)	75.1 (18.9)
Vitality	59.6 (19.9)	48.5 (26.1)
Social functioning	87.9 (20.2)	86.7 (22.2)
Emotional role functioning	88.9 (26.9)	87.7 (29.2)
Mental health	81.0 (14.3)	77.3 (18.6)

z-scores and features of score distributions showed approximately similar results for unweighted and oblique scores (Table 3). An exception was that the unweighted RAND-12 composite scores had slightly higher ceiling values, although none of them were above $\geq 15\%$. The corrected item-total correlations for the unweighted RAND-36/12 composite scores ranged from 0.52 to 0.74, indicating scores that consist of highly correlated variables. The correlations between the RAND-36/12 composite scores and the respective subscale scores showed a pattern with oblique PCS scores having stronger correlations with the subscales representing mental health and vice versa (Tables 4–5). The correlations between the composite scores derived from the two methods were very strong ($r = 0.97$ to 0.99) for both the RAND-36 and RAND-12 (Table 6). The correlations between the PCS and MCS derived by the two methods were weaker for the unweighted method than for the oblique method for both the RAND-36 ($r = 0.61$ vs. $r = 0.78$) and RAND-12 ($r = 0.58$ vs. $r = 0.79$). The effect sizes of the associations between the oblique versus unweighted composite scores and other variables had comparable magnitudes, indicating similar convergent validity (Table 7).

Discussion

We found strong correlations between the composite scores derived by the two methods for both the RAND-36 and RAND-12 and that the effect sizes of the associations between the oblique versus the unweighted composite scores and other variables had comparable magnitudes, also indicating similar convergent validity. The features of score distributions of the corresponding composite scores showed approximately similar results, except for unweighted RAND-12 composite scores having slightly higher ceiling effects.

To the best of our knowledge, this is the first study to report both the criterion validity and convergent validity of unweighted RAND-36/12 composite scores. However, two prior studies have reported the criterion validity of the RAND-36 or RAND-12 composite scores using two other methods for constructing unweighted scores. Grassi et al. [30] used data from the European Community Respiratory Health Survey and compared SF-36 composite scores derived from oblique PCA with those from an unweighted scoring system. The unweighted PCS was calculated as the sum of 18 items, while the MCS included 19 items. The correlation between the oblique and unweighted PCS was 0.97, and 0.96 between oblique and unweighted MCS. The correlation between the unweighted PCS and MCS was 0.61.

Hagell et al. [5] applied data from people with Parkinson's disease and stroke to compare SF-12 composite scores derived from the RAND-12 HSI algorithm

Table 3 Descriptive statistics of the RAND-36/12 composite z-scores (means = 0 and standard deviations = 1) and features of score distributions

RAND-36 subscales	RAND-36 PCS unweighted	RAND-36 PCS oblique	RAND-36 MCS unweighted	RAND-36 MCS oblique	RAND-12 PCS unweighted	RAND-12 PCS oblique	RAND-12 MCS unweighted	RAND-12 MCS oblique
Median	0.40	0.37	0.33	0.25	0.55	0.37	0.28	0.23
Mode	1.00	1.20	0.62	1.53	0.82	0.82	0.67	0.78
Skewness	- 1.10	- 1.10	- 1.63	- 1.21	- 1.09	- 1.11	- 1.24	- 1.08
Kurtosis	0.20	0.38	2.60	1.59	0.15	0.40	1.28	1.11
% Floor	0	0	0	0	0.1	0.1	0.1	0.1
% Ceiling	2.7	0.7	1.5	0.7	8.3	1.3	3.9	1.3
Corrected item-total correlation	PF = 0.70 RP = 0.74 BP = 0.69 GH = 0.68	NA	VT = 0.63 SF = 0.71 RE = 0.58 MH = 0.72	NA	PF = 0.66 RP = 0.72 BP = 0.71 GH = 0.61	NA	VT = 0.58 SF = 0.62 RE = 0.52 MH = 0.68	NA

PCS Physical Composite Summary; MCS Mental Composite Summary; PF physical functioning, RP physical role functioning, BP bodily pain, GH general health, VT vitality, SF social functioning, RE role functioning, MH mental health, NA not applicable

Table 4 Pearson correlations between RAND-36 subscales and unweighted and oblique composite scores

RAND-36 subscales	PCS unweighted	PCS oblique	MCS unweighted	MCS oblique
Physical functioning	0.80	0.77	0.44	0.47
Physical role functioning	0.90	0.88	0.52	0.56
Bodily pain	0.83	0.82	0.49	0.57
General health	0.80	0.82	0.61	0.71
Vitality	0.61	0.72	0.79	0.88
Social functioning	0.60	0.69	0.85	0.80
Emotional role functioning	0.42	0.48	0.82	0.70
Mental health	0.38	0.46	0.82	0.86

PCS Physical Composite Summary; MCS Mental Composite Summary

Table 5 Spearman rank correlations between the RAND-12 subscales and unweighted and oblique RAND-12 composite scores

RAND-12 subscales	PCS unweighted	PCS oblique	MCS unweighted	MCS oblique
Physical functioning	0.72	0.66	0.40	0.43
Physical role functioning	0.81	0.78	0.48	0.53
Bodily pain	0.79	0.79	0.49	0.56
General health	0.82	0.75	0.52	0.59
Vitality	0.55	0.73	0.86	0.86
Social functioning	0.53	0.63	0.73	0.68
Emotional role functioning	0.38	0.44	0.62	0.56
Mental health	0.32	0.50	0.79	0.83

PCS Physical Composite Summary; MCS Mental Composite Summary

that produced similar results to scores based on oblique PCA. The unweighted PCS was calculated as the raw sum of six items, while the MCS was from six other distinct items. The correlation between the weighted and unweighted scores were 0.99 for both PCS and MCS. The correlation between the unweighted PCS and MCS was 0.68.

The scoring methods in these two studies differed slightly from ours by using the sum of items to create raw scores, while we used unweighted linear combinations of subscale scores based on items that were standardized, ranging from 0 to 100. We think that a two-step method that initially scores the subscales and then uses them to create composite scores is more intuitive, considering

Table 6 Pearson correlations between the RAND-36/12 composite scores

	RAND-36 PCS unweighted	RAND-36 PCS oblique	RAND-12 PCS unweighted	RAND-12 PCS oblique	RAND-36 MCS unweighted	RAND-36 MCS oblique	RAND-12 MCS unweighted	RAND-12 MCS oblique
RAND-36 PCS unweighted	1							
RAND-36 PCS oblique	0.99	1						
RAND-12 PCS unweighted	0.97	0.96	1					
RAND-12 PCS oblique	0.96	0.98	0.98	1				
RAND-36 MCS unweighted	0.61	0.71	0.60	0.73	1			
RAND-36 MCS oblique	0.68	0.78	0.66	0.79	0.97	1		
RAND-12 MCS unweighted	0.59	0.69	0.58	0.72	0.97	0.96	1	
RAND-12 MCS oblique	0.66	0.75	0.66	0.79	0.94	0.97	0.97	1

Correlations between scores reflecting the same construct are marked in bold

PCS Physical Composite Summary; MCS Mental Composite Summary

Table 7 Spearman rank correlations between RAND-36/12 composite scores and related variables

	Age	Sex	Body mass index	Physical Activity	Rheumatic disease	Depression
RAND-36 PCS unweighted	-0.21	0.06	-0.21	0.27	-0.26	-0.21
RAND-36 PCS oblique	-0.13	0.08	-0.20	0.27	-0.26	-0.24
RAND-12 PCS unweighted	-0.25	0.06	-0.23	0.30	-0.24	-0.20
RAND-12 PCS oblique	-0.12	0.10	-0.19	0.29	-0.23	-0.25
RAND-36 MCS unweighted	0.11	0.10	-0.07	0.20	-0.14	-0.33
RAND-36 MCS oblique	0.10	0.10	-0.09	0.22	-0.17	-0.33
RAND-12 MCS unweighted	0.10	0.11	-0.07	0.21	-0.15	-0.32
RAND-12 MCS oblique	0.09	0.11	-0.09	0.23	-0.17	-0.32

Age (years, continuous); sex (women = 0, men = 1), body mass index (units, continuous), physical activity (strenuous physical activity: never = 0, less than 1 h per week = 1, 1–2 h per week = 2, ≥ 3 h per week = 3), rheumatic disease (no = 0, yes = 1), depression (no = 0, yes = 1)

that the subscales have a different number of items. However, the practical difference between our approach and the two other unweighted approaches for scoring composite scores seems to be minor. These findings are not surprising, given the strong correlations between the items that contribute to the RAND-36/12 composite scores.

We found that the correlations between the unweighted RAND-36/12 PCS and MCS were weaker than those created from oblique PCA. A reason for this is that oblique PCA produces weights for creating PCS and MCS that increase the correlation between these scores [7]. In the unweighted approach, no restraints are imposed, and the PCS and MCS are completely free to correlate. This

could be a strength favouring unweighted RAND-36/12 composite scores, as correlations approaching 0.80 may induce multicollinearity if the PCS and MCS are used as independent variables in the same model [31].

Regarding convergent validity, the associations between the oblique versus the unweighted RAND-36/12 composite scores and other variables had comparable magnitudes. An exception was that age was more strongly correlated with the unweighted PCS scores than the oblique ones. This could reflect that the oblique PCS scores were based on all subscales being either negatively, neutral, or positively correlated with age. There also seems to be a subtle tendency for the oblique PCS and MCS to have more similar effect sizes than the

unweighted PCS and MCS. This probably reflects the stronger correlations between the oblique PCS and MCS.

The strengths of this study include a sufficiently large sample from a general population, and that convergent validity was examined. A limitation of the study is that weight, height, physical activity, rheumatic disease, and depression were assessed by self-reports. However, the included measures have been shown to have acceptable validity [32–34]. Second, our cross-sectional design did not allow us to study longitudinal changes in unweighted versus weighted composite scores. Thus, differences in the responsiveness of change should be explored in future studies.

The main implication of this study is that we can keep the calculation of the RAND-36/12 composite scores simple. This has several advantages, such as the standardization of scoring across studies and populations. In this paper, we calculated composite scores ranging from 0 to 100, but the data can easily be converted to T-scores. An advantage of the previous scoring approaches is good comparability between PCS-12 and PCS-36 and between MCS-12 and MCS-36 [2, 7]. Such comparability is not seen with our current proposed scoring method. However, similar comparability could be derived by regressing the unweighted RAND-36 PCS and MCS T-scores in separate models for the two unweighted RAND-12 composite scores. This could be explored in a future study using a cross-validated design. It should be emphasized that our findings do not imply that weighted composite scores of HRQoL are never useful or that prior studies using different oblique composite scores for the RAND-36/12 have led to erroneous results. However, we propose that weighting is likely to be redundant if analyses of composite scores show corrected item-total correlations values ≥ 0.4 . This knowledge should also be useful to consider when developing composite scores for new HRQoL instruments.

Conclusions

In conclusion, the unweighted RAND-36/12 composite scores demonstrated satisfactory validity. Consequently, the calculation of these composite scores can be kept simple when we want them to be free to correlate. Future studies should examine the external validity of our findings and the sensitivity of the change in unweighted versus weighted composite scores in different populations.

Abbreviations

CFA: Confirmatory factor analysis; HRQoL: Health-related quality of life; MCS: Mental composite summary; PCA: Principal component analysis; PCS: Physical composite summary; RAND: Research and development.

Acknowledgements

We thank Marianne Jensen Hjermsstad and coworkers (see information under "Availability of data and materials") for providing access to the data used in this study. We also want to acknowledge the time and effort devoted by the reviewers to help us to improve the quality of our work.

Author contributions

All authors contributed to the study conception, methods, and the paper outline. Andersen conducted the data analyses and wrote the first draft of the manuscript. The other authors commented and made suggestions for improvements of previous versions to the manuscript. All authors read and approved the final manuscript.

Funding

The funding was provided by Bergen Hospital Trust.

Availability of data and materials

The data is owned by a third party: Marianne Jensen Hjermsstad, Kjersti S. Grotmol and Håvard Loge (Regional Advisory Unit for Palliative Care, Dept. of Oncology, Oslo University Hospital, Norway). E-mail: mariajhj@medisin.uio.no. Data are however available from the corresponding author upon reasonable request and with permission of the third party.

Declarations

Ethics approval and consent to participate

According to Jacobsen et al. [21]: "the survey was conducted according to Norwegian regulations for surveys. The Regional Committee for Medical and Health Research Ethics (REC) South-east Norway approved the survey. Return of the questionnaires was regarded as informed consent. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study at the time of the survey". Note that before the current study was conducted, all personally identifiable information were permanently deleted from the data set, so that the people whom the data describe are anonymous.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interest.

Author details

¹Department of Research and Development, Centre on Patient-Reported Outcomes, Bergen Hospital Trust, Bergen, Norway. ²Department of Health and Caring Sciences, Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Bergen/Førde, Norway. ³Regional Centre for Child and Youth Mental Health and Child Welfare, NORCE, Norwegian Research Centre, Bergen, Norway. ⁴Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway. ⁵Department of Clinical Science, University of Bergen, Bergen, Norway. ⁶Department of Heart Disease, Haukeland University Hospital, Bergen, Norway. ⁷Førde Hospital Trust, Førde, Norway.

Received: 7 October 2021 Accepted: 17 May 2022

Published online: 03 June 2022

References

1. Ware JE. The SF-12v2TM how to score version 2 of the SF-12® health survey:(with a supplement documenting version 1). Quality metric; 2002.
2. Ware JE, Kosinski M. SF-36 physical and mental health summary scales: a manual for users of version 1. 1st ed. Lincoln, RI: QualityMetric Inc.; 2001.
3. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med*. 2001;33:350–7.

4. Ware JE, Kosinski M, Gandek B. SF-36 health survey : manual & interpretation guide. 2nd ed. Lincoln, RI: QualityMetric Inc; 2000.
5. Hagell P, Westergren A, Arestedt K. Beware of the origin of numbers: Standard scoring of the SF-12 and SF-36 summary measures distorts measurement and score interpretations. *Res Nurs Health*. 2017;40:378–86.
6. Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in orthopaedics: a brief guide. *J Bone Joint Surg Am*. 2015;97:1628–34.
7. Farivar SS, Cunningham WE, Hays RD. 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.
8. Ware JE, Kosinski M. Interpreting SF-36 summary health measures: a response. *Qual Life Res*. 2001;10:405–13.
9. Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? *Qual Life Res*. 2001;10:395–404.
10. Nortvedt MW, Riise T, Myhr KM, Nyland HI. Performance of the SF-36, SF-12, and RAND-36 summary scales in a multiple sclerosis population. *Med Care*. 2000;38:1022–8.
11. Ware J, Keller SD, Kosinski M. How to score the SF-12 physical and mental health summary scales. Boston: Health Institute, New England Medical Center; 1995.
12. Taft C. Vidareutveckling av RAND-36 hälsoenkät: summaindex och kortversion. Regist VGR Gothenbg. 2016.
13. Lee N, Cadogan JW. Problems with formative and higher-order reflective variables. *J Bus Res*. 2013;66:242–7.
14. Tucker G, Adams R, Wilson D. Observed agreement problems between sub-scales and summary components of the SF-36 version 2-an alternative scoring method can correct the problem. *Plos One*. 2013;8:4.
15. Fleishman JA, Selim AJ, Kazis LE. Deriving SF-12v2 physical and mental health summary scores: a comparison of different scoring algorithms. *Qual Life Res*. 2010;19:231–41.
16. Willoughby M, Holochwost SJ, Blanton ZE, Blair CB. Executive functions: formative versus reflective measurement. *Meas Interdiscip Res Perspect*. 2014. <https://doi.org/10.1080/15366367.2014.929453>.
17. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment - can we keep it simple. *J R Stat Soc Ser Stat Soc*. 1992;155:353–93.
18. Grassi M, Nucera A, Zanolin E, Omenaas E, Anto JM, Leynaert B. European community respiratory health study quality of life working G: performance comparison of Likert and binary formats of SF-36 version 1.6 across ECRHS II adults populations. *Value Health*. 2007;10:478–88.
19. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. USA: Oxford University Press; 2015.
20. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32:40–66.
21. Carlson KD, Herdman AO. Understanding the impact of convergent validity on research results. *Organ Res Methods*. 2012;15:17–32.
22. Jacobsen EL, Bye A, Aass N, Fossa SD, Grotmol KS, Kaasa S, Loge JH, Moum T, Hjermstad MJ. Norwegian reference values for the short-form health survey 36: development over time. *Qual Life Res*. 2018;27:1201–12.
23. Garin O. Ceiling effect. In: Michalos AC, editor. *Encyclopedia of Quality of life and well-being research*. Dordrecht: Springer, Netherlands; 2014. p. 631–3.
24. West SG, Finch JF, Curran PJ. Structural equation models with nonnormal variables: problems and remedies; 1995.
25. Kolotkin RL, Andersen JR. A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life. *Clin Obes*. 2017;7:273–89.
26. Andersen JR, Aasprang A, Bergsholm P, Sletteskog N, Vage V, Natvig GK. Predictors for health-related quality of life in patients accepted for bariatric surgery. *Surg Obes Relat Dis*. 2009;5:329–33.
27. Salaffi F, Di Carlo M, Carotti M, Farah S, Ciapetti A, Gutierrez M. The impact of different rheumatic diseases on health-related quality of life: a comparison with a selected sample of healthy individuals using SF-36 questionnaire, EQ-5D and SF-6D utility values. *Acta Biomed*. 2019;89:541–57.
28. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. *Prev Med*. 2007;45:401–15.
29. Funder DC, Ozer DJ. Evaluating effect size in psychological research: Sense and nonsense (vol 2, pg 156, 2019). *Adv Methods Pract Psychol Sci*. 2019;2:156.
30. Grassi M, Nucera A. European community respiratory health study quality of life working G: dimensionality and summary measures of the SF-36 v1.6: comparison of scale- and item-based approach across ECRHS II adults population. *Value Health*. 2010;13:469–78.
31. Van Steen K, Curran D, Kramer J, Molenberghs G, Van Vreckem A, Bottomley A, Sylvester R. Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. *Statist Med*. 2002;21:3865–84.
32. Kopperstad O, Skogen JC, Sivertsen B, Tell GS, Saether SMM. Physical activity is independently associated with reduced mortality: 15-years follow-up of the Hordaland health study (HUSK). *Plos One*. 2017;12:e0172932.
33. Yu E, Ley SH, Manson JE, Willett W, Satija A, Hu FB, Stokes A. Weight history and all-cause and cause-specific mortality in three prospective cohort studies. *Ann of Intern Med*. 2017;166:613.
34. Bonsaksen T, Grimholt TK, Skogstad L, Lerdal A, Ekeberg O, Heir T, Schou-Bredal I. Self-diagnosed depression in the Norwegian general population - associations with neuroticism, extraversion, optimism, and general self-efficacy. *Bmc Public Health*. 2018. <https://doi.org/10.1186/s12889-018-5990-8>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

