

# Normative data of the SF-12 health survey with validation using postmyocardial infarction patients in the Dutch population

Floortje Mols · Aline J. Pelle · Nina Kupper

Accepted: 2 February 2009 / Published online: 26 February 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

## Abstract

**Objective** To gain Dutch population norms for the Short Form-12 (SF-12), a generic health status questionnaire, in a random sample of the general population and to validate these in postmyocardial infarction (MI) patients.

**Methods** 2,301 respondents from the general population and 459 post-MI patients completed the Short Form-36 (SF-36), which was used to calculate SF-12 scores.

**Results** The SF-12 summary scores correlated highly with SF-36 summary scores, demonstrating that these scores explain the same amount of variance in health status. Significant sex differences ( $P < .001$ ) existed for both the physical component summary (PCS) and the mental component summary (MCS). Multivariate analysis of variance showed a main effect of age in oblique (PCS-12:  $P < .001$ ; MCS-12:  $P < .001$ ) and orthogonally rotated PCS scores (PCS-12\_uc:  $P < .001$ ; MCS-12\_uc:  $P = .07$ ). As expected, post-MI patients reported statistically significant and clinically relevant poorer mental ( $P < .001$ ) and physical functioning ( $P < .001$ ). Differences were less pronounced for MCS and PCS derived from orthogonal rotation data. When controlling for covariates, MI did not significantly affect PCS-12\_uc anymore in orthogonally rotated data, while PCS-12\_uc was affected by fewer covariates compared with PCS-12.

**Conclusions** This study presents Dutch population norms for the SF-12 in a large random population sample obtained from both oblique and orthogonal PCA rotation methods,

revealing systematic differences between the results based on these two methods. Furthermore, this study demonstrates the discriminative validity of the SF-12 by showing that post-MI patients differ significantly from the normative population on PCS-12 scores.

**Keywords** SF-12 · Normative data · Myocardial infarction · Health status · Rotation method

## Abbreviations

MI	Myocardial infarction
PCS	Physical component summary
MCS	Mental component summary
SF-36	Short Form-36
SF-12	Short Form-12
COPD	Chronic obstructive pulmonary disease
BMI	Body mass index
CHF	Chronic heart failure
PAD	Peripheral arterial disease
PCI	Percutaneous coronary intervention
CABG	Coronary artery bypass graft
PCA	Principle components analysis

## Introduction

Health status is an important outcome in medical care, both from the patient's and health care provider's perspective, and is often measured with the Short Form-36 (SF-36) questionnaire [1–4]. However, an important concern in psychological assessment of clinical populations is to minimize response burden of patients, stimulating the pursuit of developing shorter questionnaires. A shorter and thus more user-friendly alternative to the SF-36 is the Short

F. Mols (✉) · A. J. Pelle · N. Kupper  
CoRPS-Center of Research on Psychology in Somatic diseases,  
Department of Medical Psychology, Tilburg University,  
Warandelaan 2, PO Box 90153, 5000 LE Tilburg,  
The Netherlands  
e-mail: F.Mols@uvt.nl

Form-12 (SF-12) [5]. The SF-12 measures physical and mental health by means of two summary scores; a physical component summary (PCS) and mental component summary (MCS) [6].

The SF-12 can be employed in multiple ways, i.e., SF-12 is often used to compare health status between two groups of patients, to identify predictors of health status, and to determine health status in a specific disease population. The predicament with both the SF-36 and SF-12 component summary scores is that they are somewhat difficult to interpret, because of the weighting of items to calculate PCS and MCS [5, 6]. Normative data could facilitate the interpretation of results, because these can be used to determine whether groups or individuals score above or below average for their nationality, age or sex [7]. While SF-36 normative data, according to age and sex, are available for the Dutch population [8], normative data for the SF-12 are not. Therefore, the goal of this study was to gain nationality-specific population norms for the SF-12, according to age and sex, in a large random sample of the Dutch population. In order to further validate the SF-12 (discriminative validity) and demonstrate the relevance and possible application of these Dutch norm scores, we compared health status of postmyocardial infarction (MI) patients with our newly calculated Dutch normative scores.

## Methods

### Setting and participants

The sample comprised a random selection of 2,301 adults from the general Dutch population residing in the Southern provinces of The Netherlands (population of approximately 4 million). Quota sampling was applied to ensure that different age and sex groups were equally represented in the sample. This meant that equal numbers of men and women in different age groups were sampled (e.g.,  $N$  of men aged 30–39 years =  $N$  of women aged 60–69 years). Research assistants were responsible for distributing the questionnaires and were instructed to collect an equal number of questionnaires from each age and sex subcohort, without further specification of educational or income level. Participants were approached personally or by phone. After explaining the purpose of the study, participants received an informed consent form and a questionnaire, which were returned to the research assistants in closed envelopes. The questionnaires were entered into the database by others, guaranteeing anonymity. Returned questionnaires did not contain any explicit identifiers (i.e., names) but rather, were coded by number for purposes of data collection tracking. Approval for this

study was obtained from a local ethics committee (protocol number 2006/1101).

Additionally, data was used from a prospective follow-up study in patients recovering from MI. Patients hospitalized for acute MI ( $n = 459$ ) were recruited between May 2003 and May 2006 from four teaching hospitals (Catharina Hospital, Eindhoven; St. Elisabeth Hospital, Tilburg; TweeSteden Hospital, Tilburg; and St. Anna Hospital, Geldrop) in the Southern provinces of The Netherlands. Inclusion criteria were age  $>30$  years and hospitalization due to acute MI. Criteria for diagnosis of MI were troponin I levels more than twice the upper limit, with typical ischemic symptoms (e.g., chest pain) lasting for more than 10 min or electrocardiogram (ECG) evidence of ST segment elevation or new pathological Q-waves. For patients without typical angina, the day of MI onset was identified as the day during hospitalization with peak troponin I levels  $>1.0$  and ECG evidence of ST segment elevation or new pathological Q-waves. Exclusion criteria were significant cognitive impairments (e.g., dementia) and severe medical comorbidities that increased the likelihood of early death, such as malignant cancer. All participants were approached to participate on a voluntary basis, and could withdraw from the study at any moment without implications for future treatment. The study protocol was approved by the institutional review boards of the participating hospitals (protocol number M03/1302). Written consent was obtained from all study participants.

### Measures

#### *Sociodemographic and clinical characteristics in a random sample of the Dutch population*

Demographic variables included age, sex, marital status, and classified educational level. Clinical variables were obtained from the questionnaires as well using purpose-designed questions, and included smoking, height, weight, and illnesses for which participants had received an official diagnosis, including hypertension, diabetes mellitus, renal disease, chronic obstructive pulmonary disease (COPD), hyperlipidemia, and cardiovascular disease.

#### *Sociodemographic and clinical characteristics in post-MI patients*

Demographic variables included age, sex, marital status, and educational level. Clinical variables associated with post-MI prognosis were obtained from the patients' medical records. These included cardiac treatment [prior percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG)], body mass index (BMI), hyperlipidemia, diabetes mellitus, renal insufficiency,

COPD, chronic heart failure (CHF), peripheral arterial disease (PAD), history of hypertension, and current smoking status (self-report).

#### *Clinical significant depression and anxiety*

People are, in general, very capable in determining the presence of depression, as a single question (e.g., “Are you depressed?”) was found to predict the presence of major depressive disorder very well [9]. Likewise, we obtained information on clinically diagnosed anxiety and depression in the general Dutch population by the following question in our questionnaire; “Did a doctor/medical specialist diagnose you with one of the following conditions?” This was followed by a list that included both anxiety disorder as well as depression.

In post-MI patients, the World Health Organization-authorized Dutch version of the Composite International Diagnostic Interview (CIDI) [10, 11], a fully structured diagnostic interview, was used 3 months post-MI, to determine lifetime diagnoses of major depressive disorder and anxiety disorder (consisting of panic disorder, social phobia, and/or generalized anxiety disorder). These disorders were assessed using the definitions and criteria of Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [12]. The CIDI has acceptable interrater and test–retest reliability for most nonpsychotic diagnoses, including major depressive disorder [13, 14].

#### *Health status*

In order to provide a shorter alternative to the SF-36, the SF-12 health survey was developed in 1994 [15] in the USA, and was purposely designed for large-scale measurements for which the SF-36 was too lengthy. The SF-12 contains 12 items derived from the SF-36, including one or two items from each of the eight SF-36 subscales (physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health) [8]. These 12 items are used to construct the physical component summary (PCS) and the mental component summary (MCS) [15].

In this study, health status was assessed by version 1 of the Dutch SF-36 questionnaire (in both the norm and post-MI population), which was used to calculate SF-12 component summary scores [5]. Internal consistency reliability coefficients have been reported for each SF-36 subscale, ranging from 0.62 to 0.96 with a median of 0.80 [16]. Test–retest reliability of the SF-36 ranges from 0.43 to 0.90 with a median of 0.64 after 6 months in the general population [16]. In post-MI patients, health status was assessed 3 months post-MI by means of the SF-36.

## **Statistical analysis**

### *Calculating SF-36 domain scores*

Following the SF-36 scoring instructions, items were reversed or recalibrated when necessary [17]. Then, the raw domain scores were calculated, also including those respondents who had missing values on no more than half of the domain’s score items. In a third step, the raw domain scores were transformed to a 0–100 scale. For those respondents who had missing values on more than half of the domain’s score items, no data substitution algorithms were used to handle missing data. For the eight dimensions of the SF-36, 0.5% ( $n = 12$ ) of cases were missing for physical functioning, 0.7% ( $n = 12$ ) of cases were missing for social functioning, 1.8% ( $n = 42$ ) of cases were missing for role limitations due to physical health, 2.1% ( $n = 48$ ) of cases were missing for role limitations due to emotional problems, 0.9% ( $n = 20$ ) of cases were missing for mental health, 0.8% (19) of cases were missing for vitality, 0.7% ( $n = 17$ ) of cases were missing for bodily pain, and 1.2% (27) of cases were missing for general health perception.

### *Calculating PCS-36 and MCS-36 component summary scores*

PCS and MCS scores are essentially summations of the weighted domain scores. To obtain PCS and MCS scores, a dedicated procedure was used, described in the SF-36 summary scales manual [18]. The domain scores were first standardized, and then principle components analysis (PCA) was employed to obtain the domain weights needed to construct PCS and MCS. More specifically, the eight domains were factor-analyzed twice. Once using the standardized approach, orthogonal rotation (varimax), which by definition assumes factors to be uncorrelated, and once using an approach, oblique rotation (promax), which by definition allows factors to be correlated. The SF-36 summary measures manual [18] and various papers show that, when using the orthogonal method, the mental and physical component summaries demonstrate low to very low empirical correlations [19, 20]. Papers that have used the oblique rotation method have reported high correlations between the mental and physical component summary scores [15, 21–23]. The reason for employing orthogonal rotation, as stated in the SF-36 manual, was a more straightforward interpretation of each component, driven by the goal of testing the construct validity of the SF-36. The use of oblique rotation in this and previous papers on the other hand had different reasons, in that it was based on theoretical arguments about the a priori knowledge of the relation between mental and physical health. We decided

therefore to analyze the data twice because of the apparent discrepancy between the manual guidelines advocating orthogonal rotation for their purposes, and the observations from the literature that mental and physical health in fact are related [15, 21–23]. The domain weights (resulting from orthogonal and oblique rotation of the eight SF-36 subscales) are presented in Table 1, where a comparison is made between the new Dutch weights and the gold-standard US weights (adapted from [24]). Note that the US weights were obtained from PCA using orthogonal rotation (varimax). The observable differences in Table 1 are largely due to the difference in rotation. When we applied an orthogonal rotation, despite factors being correlated, our Dutch weights resembled US weights closely. Additionally, the negative loadings largely disappear in case of oblique rotation, removing the negative effect of better mental health on physical health and vice versa, as Farivar et al. suggested [25].

The final step was to calculate raw PCS-36 and MCS-36 component scores by summing up all weighted domain scores according to the weights in the two-factor solution based on orthogonal and oblique rotation. Summary scores were then transformed so that they would have standard deviation of 10 (multiply by 10) and mean of 50 (add 50).

#### Calculating PCS-12 and MCS-12 component summary scores

In accordance with the developers' recommendations [6] we selected the appropriate SF-12 items from the SF-36 to reproduce the PCS-12 and MCS-12 scores in the general Dutch population and a large sample of Dutch post-MI patients. Instructions were followed to reverse item scores so that a higher score always represented better health

status. Next, indicator variables were made for each answer option, and PCS-12 and MCS-12 indicator variables weights were calculated by regressing the indicator variables against, respectively, the PCS-36 and MCS-36 component summary scores. Table 2 shows the indicator variable weights for this Dutch normative population. These weights should be used to calculate SF-12 component summary scores when comparing data with this normative population. Again, we calculated the weights for each rotation method. When comparing the obtained weights with the weights published by Farivar et al. based on US data, the magnitude of the weights is comparable [25].

Rotation-specific PCS-12 and MCS-12 component summary scores were finally computed by summing all weighted indicator variable scores, and standardizing them by adding the constants (specific for PCS-12 and MCS-12, respectively) from the regression analyses described above. For the validation sample of Dutch post-MI patients we used our normative population weights and constants in the calculation of PCS-12 and MCS-12 scores. When presenting our results, the extension “\_uc” (which stands for *uncorrelated*) will indicate that summary scores were derived from orthogonal rotation PCA.

#### Statistical analyses

Correspondence of SF-12 scores with the SF-36 scores was checked by examining the proportion of the variance in PCS-36 and MCS-36 component summary scores that was explained by the 12 items of the SF-12 ( $R^2$ ). In addition, we examined the (Spearman's) correlations between the summary component scores of the SF-36 and SF-12, and between the summary component scores within the SF-12. In addition, we correlated (Spearman's) SF-12 summary

**Table 1** Factor score coefficients from the SF-36 resulting from PCA using two rotation methods

	PCS-36			MCS-36		
	The Netherlands		USA	The Netherlands		USA
	Oblique	Orthogonal	Orthogonal	Oblique	Orthogonal	Orthogonal
Physical functioning	0.328	0.411	0.424	−0.042	−0.216	−0.230
Role physical	0.244	0.258	0.351	0.079	−0.023	−0.123
Bodily pain	0.312	0.369	0.318	−0.081	−0.169	−0.097
General health perception	0.276	0.291	0.250	0.028	−0.073	−0.016
Energy/vitality	0.095	0.000	0.029	0.265	0.283	0.235
Social functioning	0.113	0.044	−0.008	0.238	0.229	0.269
Role emotional	−0.054	−0.193	−0.192	0.341	0.431	0.434
Mental health	−0.069	−0.220	−0.221	0.384	0.490	0.486

*Note:* US weights are the result of PCA with orthogonal rotation and those from the US SF-36 developers (adapted from [17, 24]). Dutch factor score coefficients were obtained from two PCA analyses, one employing orthogonal rotation (standardized method) and one using oblique rotation

**Table 2** Regression weights for calculating SF-12 PCS and MCS scores

SF-12 answer category	Oblique rotation method		Orthogonal rotation method	
	PCS-12	MCS-12	PCS-12_uc	MCS-12_uc
<i>Physical functioning</i>				
PFQ2_A1	−4.719	0.879	−6.148	3.072
PFQ2_A2	−2.352	0.306	−2.987	1.334
PFQ4_A1	−6.461	0.459	−7.757	3.016
PFQ4_A2	−2.755	0.142	−3.346	1.338
<i>Physical role</i>				
rpQ2_A1	−2.858	−0.863	−3.041	0.070
rpQ3_A1	−3.911	−1.035	−4.246	0.407
<i>Emotional role</i>				
REQ2_A1	0.594	−5.599	3.016	−6.910
REQ3_A1	0.760	−6.043	3.336	−7.438
<i>Bodily pain</i>				
BPQ2_A1	−12.066	0.371	−14.346	4.565
BPQ2_A2	−10.585	1.787	−12.952	5.433
BPQ2_A3	−7.174	1.432	−8.835	3.850
BPQ2_A4	−3.998	0.558	−4.926	2.083
<i>General health</i>				
GHQ1_A1	−9.956	−2.736	−10.904	0.756
GHQ1_A2	−6.679	−0.868	−7.642	1.556
GHQ1_A3	−3.447	−0.318	−3.982	0.928
GHQ1_A4	−1.665	−0.126	−1.958	0.508
<i>Vitality</i>				
VTQ2_A1	−1.895	−6.232	0.307	−6.640
VTQ2_A2	−2.652	−6.333	−0.657	−6.441
VTQ2_A3	−1.907	−4.611	−0.420	−4.775
VTQ2_A4	−1.238	−3.303	−0.168	−3.407
VTQ2_A5	−0.755	−1.645	−0.208	−1.705
<i>Mental health</i>				
MHQ3_A1	0.341	−7.603	2.966	−8.817
MHQ3_A2	0.711	−9.250	4.149	−10.865
MHQ3_A3	0.189	−7.191	2.801	−8.272
MHQ3_A4	0.075	−4.234	1.598	−4.858
MHQ3_A5	−0.085	−2.119	0.676	−2.439
MHQ4_A1	2.695	−11.164	6.535	−13.668
MHQ4_A2	0.525	−11.018	4.526	−12.763
MHQ4_A3	0.007	−8.347	2.932	−9.513
MHQ4_A4	0.492	−4.847	2.231	−5.650
MHQ4_A5	0.119	−2.325	0.920	−2.632
<i>Social functioning</i>				
SFQ2_A1	−4.560	−7.505	−1.885	−7.110
SFQ2_A2	−3.387	−8.255	−0.464	−8.331
SFQ2_A3	−2.866	−4.661	−1.368	−4.328
SFQ2_A4	−1.316	−2.348	−0.495	−2.346

*Note:* The PCS and MCS component scores are calculated by first adding all weighted answer categories, and then standardizing this score by adding the regression constant (oblique: PCS constant = 61.14074, MCS constant = 61.63188; orthogonal: PCS constant = 58.752, MCS constant = 59.289). For each question, indicator variables were constructed based on answer categories (see “Methods” section). Each row represents the answer categories per subscale. For example “PFQ2\_A1” stands for answer option 1 on question 2 of the physical functioning subscale

component scores (calculated once for each rotation method) with demographic characteristics, i.e., age, sex, educational level, and social situation.

Data on patient and disease characteristics were compared between the Dutch normative population and the post-MI patient population using independent samples

*t*-tests for continuous variables and chi-squared analyses for categorical variables.

Mean SF-12 component summary scores were calculated for the Dutch normative population, separately for each rotation method (Tables 3 and 4). Norm scores were categorized by sex (total group; male; female) and by age (total group; 30–39 years; 40–49 years; 50–59 years; 60–69 years; 70–79 years). Sex and age differences were formally tested using the nonparametric Mann–Whitney *U*-test (sex), and a multivariate analysis of variance (MANOVA) with a Dunnett's T3 post hoc analysis for unequal variances (categorized age).

To evaluate the discriminative properties of the Dutch SF-12, MANOVA was used to examine the differences in SF-12 component summary scores between the normative and post-MI patient samples. Analyses were performed for the total group and for each age category separately, for both variants of the summary scores. To evaluate clinical relevance of the observed differences in health status scores, effect sizes were calculated using Cohen's *d* [26, 27]. Cohen's *d* represents the differences between means

(i.e., normative sample versus post-MI patients) divided by the pooled standard deviation. An effect size ranging from 0.00 to 0.20 is considered negligible to small, from 0.20 to 0.50 small to moderate, from 0.50 to 0.80 large, and >0.80 very large [27]. In a next step, demographics and clinical and psychological risk factors (see Table 5 for an overview of these variables) were added as covariates to examine the extent to which the differences between the normative population and the post-MI population might be explained by these covariates. For this purpose, we dichotomized educational level into low (high school and below) versus higher education. All statistical analyses were performed using SPSS version 14.0. A *P*-value < .05 was used for all tests to indicate statistical significance.

## Results

The sample from the general Dutch population comprised 2,301 middle-aged adults (49.1% men; age: mean 55.2 years, SD 14.3 years). Descriptives are presented in Table 5.

**Table 3** Dutch mean age- and sex-standardized scores for the mental and physical component summary of the SF-12 (oblique rotation method)

Age (years)	Total group ( <i>n</i> = 2,013)		Women ( <i>n</i> = 1,019)		Men ( <i>n</i> = 992)	
	MCS-12	PCS-12	MCS-12	PCS-12	MCS-12	PCS-12
30–39	<i>n</i> = 395	<i>n</i> = 395	<i>n</i> = 198	<i>n</i> = 198	<i>n</i> = 197	<i>n</i> = 197
Mean score	51.08	53.77	49.80	52.76	52.36	54.78
SD	9.22	6.44	9.61	6.71	8.65	6.02
95% CI	50.17–51.99	53.13–54.4	48.46–51.15	51.82–53.70	51.14–53.57	53.93–55.63
40–49	<i>n</i> = 421	<i>n</i> = 421	<i>n</i> = 222	<i>n</i> = 222	<i>n</i> = 198	<i>n</i> = 198
Mean score	51.07	53.17	49.79	51.97	52.57	54.55
SD	9.09	6.99	9.58	7.48	8.25	6.13
95% CI	50.20–51.94	52.50–53.84	48.52–51.06	50.98–52.96	51.41–53.73	53.69–55.41
50–59	<i>n</i> = 436	<i>n</i> = 436	<i>n</i> = 216	<i>n</i> = 216	<i>n</i> = 220	<i>n</i> = 220
Mean score	50.89	50.71	49.48	49.58	52.27	51.83
SD	8.85	9.07	9.41	9.66	8.04	8.32
95% CI	50.06–51.72	49.86–51.57	48.22–50.74	48.28–50.87	51.20–53.34	50.72–52.93
60–69	<i>n</i> = 391	<i>n</i> = 391	<i>n</i> = 194	<i>n</i> = 194	<i>n</i> = 197	<i>n</i> = 197
Mean score	51.30	50.86	50.07	49.76	52.50	51.95
SD	8.76	8.26	8.84	8.76	8.53	7.60
95% CI	50.43–52.17	50.04–51.68	48.82–51.32	48.52–51.00	51.31–53.70	50.88–53.02
70–79	<i>n</i> = 370	<i>n</i> = 370	<i>n</i> = 189	<i>n</i> = 189	<i>n</i> = 180	<i>n</i> = 180
Mean score	47.71	44.22	46.59	42.16	48.87	46.46
SD	10.52	11.69	10.43	11.96	10.55	10.99
95% CI	46.63–48.78	43.02–45.41	45.09–48.09	40.44–43.88	47.32–50.42	44.84–48.07
Total group	<i>n</i> = 2,013	<i>n</i> = 2,013	<i>n</i> = 1,019	<i>n</i> = 1,019	<i>n</i> = 992	<i>n</i> = 992
Mean score	50.45	50.66	49.19	49.38	51.78	52.00
SD	9.36	9.23	9.64	9.74	8.89	8.46
95% CI	50.04–50.87	50.26–51.06	48.59–49.78	48.78–49.98	51.22–52.33	51.48–52.53

*Abbreviations* SD = standard deviation, CI = confidence interval, PCS = physical component summary, MCS = mental component summary

**Table 4** Dutch mean age- and sex-standardized scores for the mental and physical component summary of the SF-12 (orthogonal rotation method)

Age (years)	Total group ( <i>n</i> = 2,013)		Women ( <i>n</i> = 1,019)		Men ( <i>n</i> = 992)	
	MCS-12_uc	PCS-12_uc	MCS-12_uc	PCS-12_uc	MCS-12_uc	PCS-12_uc
30–39	<i>n</i> = 395	<i>n</i> = 395	<i>n</i> = 198	<i>n</i> = 198	<i>n</i> = 197	<i>n</i> = 197
Mean score	49.78	54.06	48.67	53.37	50.89	54.76
SD	9.70	6.53	10.31	7.09	8.92	5.85
95% CI	48.82–50.74	53.42–54.71	47.22–50.11	52.38–54.37	49.63–52.14	53.93–55.58
40–49	<i>n</i> = 421	<i>n</i> = 421	<i>n</i> = 222	<i>n</i> = 222	<i>n</i> = 198	<i>n</i> = 198
Mean score	49.99	53.37	48.95	52.45	51.22	54.41
SD	9.18	6.68	9.78	7.31	8.30	5.77
95% CI	49.11–50.87	52.73–54.01	47.66–50.24	51.48–53.41	50.05–52.38	53.60–55.22
50–59	<i>n</i> = 436	<i>n</i> = 436	<i>n</i> = 216	<i>n</i> = 216	<i>n</i> = 220	<i>n</i> = 220
Mean score	50.72	50.51	49.51	49.71	51.91	51.31
SD	9.04	9.45	9.48	9.91	8.45	8.91
95% CI	49.87–51.57	49.63–51.40	48.24–50.78	48.28–50.87	50.78–53.03	50.12–52.49
60–69	<i>n</i> = 391	<i>n</i> = 391	<i>n</i> = 194	<i>n</i> = 194	<i>n</i> = 197	<i>n</i> = 197
Mean score	51.13	50.54	50.13	49.70	52.12	51.37
SD	8.73	8.24	9.02	9.06	8.35	7.28
95% CI	50.26–52.00	49.72–51.36	48.85–51.40	48.42–50.98	50.95–53.29	50.35–52.39
70–79	<i>n</i> = 370	<i>n</i> = 370	<i>n</i> = 189	<i>n</i> = 189	<i>n</i> = 180	<i>n</i> = 180
Mean score	49.50	44.06	48.99	42.06	50.01	46.26
SD	8.49	11.15	9.50	11.63	9.48	10.17
95% CI	46.53–50.47	42.93–45.20	47.62–50.35	40.39–43.73	48.61–51.40	44.76–47.76
Total group	<i>n</i> = 2,013	<i>n</i> = 2,013	<i>n</i> = 1,019	<i>n</i> = 1,019	<i>n</i> = 992	<i>n</i> = 992
Mean score	50.24	50.63	49.25	49.60	51.26	51.71
SD	9.24	9.21	9.63	9.88	8.71	8.30
95% CI	49.83–50.64	50.22–51.03	48.65–49.84	48.99–50.20	50.72–51.81	51.19–52.23

*Abbreviations* SD = standard deviation, CI = confidence interval, PCS = physical component summary, MCS = mental component summary

#### Confirmation of the relation between SF-12 and SF-36

As expected, results showed that there was almost complete overlap between the SF-12 and SF-36, as the proportion of the variance ( $R^2$ ) in PCS-36 and MCS-36 explained by SF-12 items was, respectively, 94% and 93% in case of oblique rotation, and 92% for both component summaries in case of orthogonal rotation. Furthermore, we found that, in case of oblique rotation, PCS-36 and PCS-12 correlated .93 and MCS-36 and MCS-12 summary scores correlated .96. Using orthogonal rotation, correlations were .90 and .94 for the relation between PCS-36 and PCS-12, and MCS-36 and MCS-12, respectively. When allowing a correlation between PCS-12 and MCS-12 (by PCA with oblique rotation), the SF-12 derived summary scores PCS-12 and MCS-12 showed substantial correlation ( $r = .61$ ), but this was no different from the SF-36 summary scores ( $r = .59$ ). Nonparametric correlations (Spearman's rho) between MCS-12 and PCS-12 and demographic variables sex, age, education level, and marital status for both rotation methods are presented in Table 6. Sex and age

differences were assessed for all component summary scores, calculated using oblique and orthogonal rotation method. The Mann–Whitney test showed the presence of significant sex differences for PCS-12 ( $P < .001$ ) and PCS-12\_uc ( $P < .001$ ) and MCS-12 ( $P < .001$ ) and MCS-12\_uc ( $P < .001$ ). MANOVA showed a significant main effect of age (PCS-12:  $F = 73.270$ ,  $P < .001$ ; MCS-12:  $F = 10.067$ ,  $P < .001$ ). Post hoc analysis showed that, for PCS-12, subjects in age groups 30–39 years and 40–49 years were similar to each other ( $P = .90$ ), and different from the other age groups ( $ps < .001$ ). Furthermore, PCS-12 scores for age groups 50–59 years and 60–69 years were similar ( $P = 1.00$ ), and the oldest participants (aged 70–79 years) differed from all other groups ( $ps < .001$ ). The mental component summary scores (MCS-12) were of equal magnitude for the first four age groups, while the oldest participants (70–79 years) scored significantly lower than all other groups ( $P < .001$ ). For PCS-12\_uc similar age differences were found ( $F = 81.308$ ,  $P < .001$ ) and post hoc analysis revealed the same pattern of similarities and differences between age groups as the norm scores based

**Table 5** Baseline characteristics for the Dutch normative population and post-MI patients

Characteristic	Normative population ( <i>n</i> = 2,301)	Post-MI patients ( <i>n</i> = 459)	<i>P</i> -value
<i>Demographics</i>			
Age, years, mean (SD)	55.2 (14.3)	59.5 (11.2)	<.001
Male sex	49.1 (1,128)	80.0 (367)	<.001
No partner	16.5 (380)	17.6 (81)	.55
Low education	30.2 (694)	45.3 (208)	<.001
<i>Cardiovascular history</i>			
Cardiovascular disease <sup>#</sup>	12.8 (289)	100 (459)	<.001
<i>Clinical factors</i>			
Diabetes mellitus	5.6 (126)	14.2 (65)	<.001
Hypercholesterolemia	1.0 (22)	34.7 (157)	<.001
Hypertension	9.3 (208)	38.7 (176)	<.001
COPD	4.8 (110)	7.8 (36)	<.01
Renal failure	0.4 (9)	3.5 (16)	<.001
Current smoking	20 (459)	38.0 (174)	<.001
BMI, kg/m <sup>2</sup> , mean (SD)	25.6 (3.9)	26.9 (4.0)	<.001
<i>Psychological risk factors</i>			
Anxiety disorder	1.8 (40)	7.1 (29)	<.001
Depressive disorder	4.0 (89)	17.1 (70)	<.001

Results are presented as % (*n*), unless otherwise stated

<sup>#</sup> Cardiovascular disease includes previous MI, angina pectoris, atherosclerosis, chronic heart failure, peripheral arterial disease or any procedure relating to the heart (coronary artery bypass graft or percutaneous coronary intervention)

BMI = body mass index; COPD = chronic obstructive pulmonary disease

**Table 6** Nonparametric correlations between MCS-12 and PCS-12 and demographic characteristics

	Oblique rotation		Orthogonal rotation	
	PCS-12	MCS-12	PCS-12 <sub>uc</sub>	MCS-12 <sub>uc</sub>
Sex	−0.15	−0.16	−0.10	−0.12
Categorized age	−0.29	−0.09	−0.33	0.01 (ns)
Education*	−0.19	−0.09	−0.18	−0.03
Marital status*	−0.13	−0.13	−0.10	−0.09

All correlations were significant at a *P* < .01 level, except age × MCS12<sub>uc</sub>. \*These variables are dichotomized, i.e., education: lower versus higher education, and marital status: having a partner versus having no partner

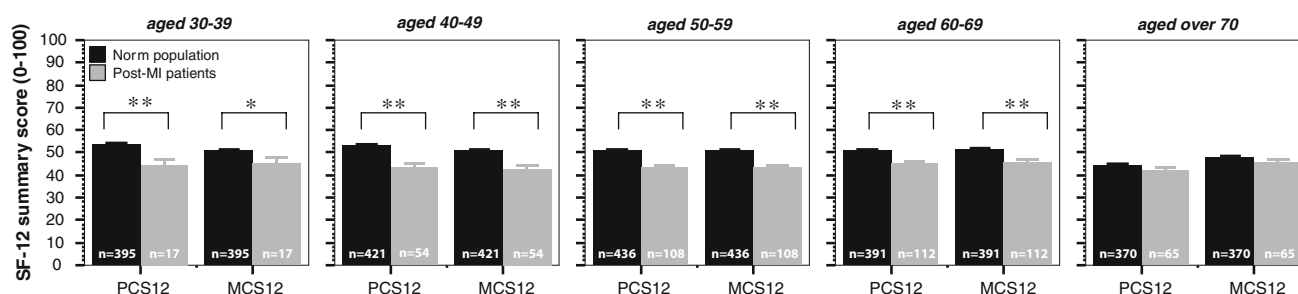
on the oblique rotation method. However, for MCS-12<sub>uc</sub> no age differences were found (*F* = 2.130, *P* = .08), with post hoc analysis showing no group differences (*ps* > .13).

Population norms for the Dutch population differed from those of Dutch post-MI patients

The sample of Dutch post-MI patients comprised 459 individuals (80% men, age: mean 59.5 years, SD 11.2 years) (Table 5). Compared with the normative Dutch sample, post-MI patients were older, more often male, and had a lower educational level (all *ps* < .001). Not surprisingly, comorbid diseases and biomedical risk factors were more prevalent in the post-MI population (all *ps* < .01), as well as having a history of anxiety disorder or depressive disorder, which was more common in post-MI patients (*ps* < .001). None of the post-MI patients had been diagnosed with CHF or PAD. Five percent of post-MI patients underwent a CABG or PCI.

Post-MI patients had statistically significant poorer scores than the general population on MCS-12 (means 50.6 (±9.3) versus 44.6 (±12.4); *F* = 105.445, *P* < .001) and PCS-12 (means 50.7(±9.2) versus 43.5(±10.7); *F* = 168.222, *P* < .001) of the SF-12 (Fig. 1). These differences were also clinically relevant, with Cohen's *d* being large for MCS-12 and PCS-12 (0.55 and 0.72, respectively) [27]. Orthogonal rotation derived summary scores essentially resulted in the same differences. However, differences in MCS-12<sub>uc</sub> (means 50.2 (±9.2) versus 46.2 (±12.0); *F* = 53.044, *P* < .001) and PCS-12<sub>uc</sub> (means 50.6 (±9.2) versus 44.6 (±10.1); *F* = 127.456, *P* < .001) between the post-MI patients and normative control group were less pronounced, with Cohen's *d* being 0.37 for MCS-12<sub>uc</sub> (i.e., small to moderate) and 0.62 (i.e., large) for PCS-12<sub>uc</sub>. When stratifying data by age categories, results demonstrated that in all age categories post-MI patients reported a lower mental (all *ps* < .01) and physical (all *ps* < .001) health status (oblique rotation), except for the oldest subgroup (aged 70–79 years) in which MCS-12 (*F* = 1.974, *P* = .16) and PCS-12 (*F* = 1.814, *P* = .18) scores were equal to the normative age group. Orthogonal rotation derived summary scores showed slightly different results for MCS-12<sub>uc</sub>, as there were no differences in mental health status between MI patients and normative controls in the youngest age cohort (*F* = 1.994, *P* = .16) or in the oldest age group (*F* = 1.241, *P* = .27). PCS-12<sub>uc</sub> was significantly impaired in all MI patients (*ps* < .001) except for those in the oldest age group (*F* = 1.320, *P* = .25).

Adding the covariates to the nonstratified multivariable model revealed that all covariates except hyperlipidemia (*P* = .14) and lower educational level (*P* = .25) significantly affected the physical component summary score (all *ps* < .01). The mental component summary score significantly covaried with all variables (all *ps* < .05) except for BMI (*P* = .43), hyperlipidemia (*P* = .20), and lower educational level (*P* = .25). After controlling for all potential confounders, MI patients still differed



**Fig. 1** Comparison of SF-12 data (oblique rotation) from the general Dutch population with a sample of post-MI patients. Black bars represent the normative population; gray bars represent the post-MI

population. Error bars: 1 SEM. \*Significant at  $P < .01$  level, \*\*Significant at  $P < .001$  level

significantly from the norm population on PCS-12 scores ( $P = .04$ ), while differences in MCS-12 scores were no longer significant ( $P = .16$ ). For the orthogonal rotation derived summary scores, quite a few differences were observed compared with the oblique rotation derived summary scores. In addition to hyperlipidemia ( $P = .21$ ) and lower educational level ( $P = .37$ ), PCS-12<sub>uc</sub> was also not affected by the covariates anxiety ( $P = .55$ ) and depression ( $P = .99$ ). The other variables significantly covaried with PCS-12<sub>uc</sub> (all  $ps < .03$ ). MCS-12<sub>uc</sub> significantly covaried with all variables (all  $ps < .05$ ) except for BMI ( $P = .20$ ), hypertension ( $P = .06$ ), hyperlipidemia ( $P = .34$ ), and lower educational level ( $P = .33$ ). After controlling for all potential confounders, data based on PCA with orthogonal rotation revealed that MI patients did no longer differ significantly from the norm population on PCS-12<sub>uc</sub> scores ( $P = .07$ ) or MCS-12<sub>uc</sub> scores ( $P = .37$ ).

## Discussion

This study confirmed the close resemblance of the SF-36 and SF-12 in the Dutch population. In line with previous studies [23], significant sex and age differences existed for both PCS-12 and MCS-12. For PCS-12, this was valid for both rotation methods. For MCS-12, norm scores (both orthogonal and oblique rotated data) did not show an age-related decline (MCS-12 scores decreased on average 3 points, but this decrease was found completely in the oldest age group), whereas this might have been expected based on a previous study in Greeks in which there was a steady decrease of in total 6 points on MCS-12 between ages 18 years and 65+ years [23]. Post-MI patients had statistically significant and clinically relevant poorer scores than the general Dutch population on MCS-12 and PCS-12. After controlling for all covariates, MI patients still differed significantly from the normative population on PCS-12 scores (oblique rotation only), while MCS-12 scores were comparable (both rotation methods).

The findings of the present study regarding the oblique rotation results concur with a study conducted in the UK in which PCS-36 and MCS-36 scores were also found to be highly correlated to the PCS-12 and MCS-12, respectively [21]. Several other studies confirmed these findings across different European countries [15, 22, 23]. The current study also showed that, within the SF-12, the MCS and PCS summary scores were substantially correlated, and this was also found for the SF-36 summary scores. Correlations between SF-12 PCS and MCS scores in the IQOLA study, an SF-12 validation study in nine European countries including The Netherlands, were quite small, which was due to the use of US PCS-36 and MCS-36, which were constructed based on PCA with orthogonal rotation, precluding a higher correlation between factors [15]. Furthermore, correlations between the two SF-12 summary scores and correlations between the SF-36 summary scores were also close to zero in a Chinese population, which is to be expected as they followed guidelines to employ orthogonal rotation [19]. Conversely, a recent study by Spindler and colleagues in a Danish sample of cardiac patients, reported a significant correlation of .33 between PCS-36 and MCS-36, while using US weights to calculate the summary scores [28]. This correlation seems intrinsic to the PCS and MCS scores, as all items, both physical and mental, contribute to both component scores.

In this study we assumed that responses to the SF-12 items embedded within the SF-36 are the same as responses to the original SF-12 items when administered alone. This assumption was tested in two studies, confirming the congruence in mean SF-12 PCS and MCS when the SF-12 was administered on its own compared with when the SF-12 items were extracted from the SF-36 [6, 29, 30].

When calculating the SF-36 PCS and MCS scores, the eight SF-36 domains were factor-analyzed using PCA to obtain the domain weights. Because we assumed the factors to be correlated, we applied oblique rotation, as is recommended in PCA [31]. After oblique rotation, factors indeed showed moderate to high correlations, confirming

our choice for oblique rotation. By using this rotation method, our domain weights differed from the gold-standard US weights, which used PCA orthogonal rotation [24]. Other studies have also used orthogonal rotation since this procedure is recommended by the developers [21, 23, 24, 32]. When we forced an orthogonal rotation, our weights resembled US weights closely. Importantly, a recent study including over 40,000 subjects from two large-scale UK population samples reported that oblique models gave the best fit to the data, and indicated a considerable correlation between PCS and MCS (between .60 and .76 depending on the adhered model) [33]. These findings warrant further detailed examination of the relation between PCS and MCS scores in previous and future studies, as well as examination of the implications for the calculation of the summary scores, and reported results.

Post-MI patients had statistically significant and clinically relevant poorer scores compared with the general Dutch population on MCS and PCS of the SF-12, and after controlling for all covariates, patients still differed significantly from the norm population on PCS scores in oblique but not in orthogonally rotated data. When controlling for covariates, MI did not significantly affect PCS-12 anymore in orthogonally rotated data, although fewer covariates significantly affected PCS (and MCS) scores.

Depression and anxiety were found to significantly explain group differences in PCS-12 scores between the general population and the post-MI population, using the oblique rotation method. This finding may be explained by the inclusion of somatic symptoms such as fatigue and sleeping problems in the definition of depression. In fact, depression in post-MI patients is confounded by cardiac health (i.e., pump function, cardiac history) [34, 35]. Therefore, it was to be expected that depression (and anxiety, a frequent comorbid condition to depression) would act as a significant covariate of the physical summary score.

To our knowledge, only one other study compared health status of MI patients, measured with the SF-12, with normative data [36]. That study also concluded that MCS and PCS scores were significantly lower in patients than in normative controls, although they did not control for possible covariates, nor did they stratify by age.

One possible explanation for these observed differences between the data obtained with the two rotation methods would be differences in SF-12 regression weights. When examining these weights (Table 2), we may observe several differences. The vitality item and the social functioning item have a stronger impact on PCS in the oblique rotation data, compared with their impact on the orthogonal rotation data. In addition, mental health items and the emotional role item load stronger on the PCS in case of orthogonal rotation, compared with the PCS

obtained from oblique rotation. When comparing the MCS scores, it is noteworthy that the bodily pain item as well as the physical functioning items load stronger on the orthogonal derived MCS. Concluding, it is clear that, in the orthogonal rotation data, the PCS score is more strongly affected by the mental health questions, compared with the oblique rotation data, while the MCS score is more strongly affected by physical health items, compared with the oblique rotation data. This would mean that, at item level, forcing the summary scores to be uncorrelated (i.e., orthogonal rotation) has the unwanted effect of cross-loading of physical and mental items on both the summary scores. This effect has been described previously [25], and now is confirmed in a population of different nationality in the current study.

Our findings demonstrated less pronounced age differences in health status scores with increasing age compared with previous studies [23, 37, 38]. The plateau in health-related quality of life may be characteristic of the Dutch population. The difference, for example, compared with US norm scores (which show a steadily decline in PCS and a stable or slightly increasing MCS as people grow older [15], especially regarding PCS), may be explained by differences in health care systems, and social security, but also by cultural differences in health appraisal. Health care in The Netherlands is available for everyone, at an acceptable cost. In addition, social security acts provide a guaranteed income for those who cannot support themselves. Therefore, becoming ill may have a smaller effect on physical health-related quality of life in The Netherlands as compared with, for example, in the USA or other European countries.

The observation in our study that health status was not related to educational level may at first seem odd, since it is known from previous studies that low socioeconomic status (SES) may negatively affect health care utilization in case patients do not have health care insurance. However, in The Netherlands everyone has equal access to care, despite SES, which can explain why a lower educational level was not significantly related to PCS and MCS scores.

Some limitations of the current study should be noted. Overall sample size of post-MI patients ( $n = 459$ ) was rather small in comparison with the Dutch norm population ( $n = 2,301$ ). Furthermore, for post-MI patients our own new Dutch weights were used, which may hamper direct comparison with other studies using different weights. However, as we presented domain weights, regression weights, and norm scores for both rotation methods, our results will be comparable with future studies that use either method. Furthermore, participants in our study were between 30 and 79 years of age, which limits possible comparability with studies outside this age range.

Despite these limitations, the current study is the first to present Dutch population norms for the SF-12, stratified by age and sex, which can be useful to interpret PCS and MCS scores from other Dutch studies using this instrument. Moreover, this study paves the way for reflection on the potential models relating PCS with MCS, and influencing PCS and MCS calculation. Finally, we showed that Dutch post-MI patients differ significantly from the normative population on PCS scores, even when controlling for disease characteristics, risk factors, and comorbid diseases. In conclusion, this study presents evidence to support the use of the SF-12 as a shorter alternative to the SF-36 in the assessment of health status in Dutch studies, particularly when overall physical and mental health are the main outcomes of interest.

**Acknowledgements** We would like to thank Dr. Elisabeth J. Martens for providing us with the data on post-MI patients. We wish to thank Dr. Frans Pouwer for his advice on some statistical procedures. Furthermore, we thank all participants for their participation in this study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Rumsfeld, J. S. (2002). Health status and clinical practice: When will they meet? *Circulation*, 106(1), 5–7. doi:10.1161/01.CIR.0000020805.31531.48.
2. Spertus, J. (2001). Selecting end points in clinical trials: What evidence do we really need to evaluate a new treatment? *American Heart Journal*, 142(5), 745–747. doi:10.1067/mhj.2001.119135.
3. Spertus, J. A., Radford, M. J., Every, N. R., Ellerbeck, E. F., Peterson, E. D., & Krumholz, H. M. (2003). Challenges and opportunities in quantifying the quality of care for acute myocardial infarction: Summary from the acute myocardial infarction working group of the American heart association/American college of cardiology first scientific forum on quality of care and outcomes research in cardiovascular disease and stroke. *Journal of the American College of Cardiology*, 41(9), 1653–1663. doi:10.1016/S0735-1097(03)00415-7.
4. Ware, J. E., Jr., & Sherbourne, C. D. (1992). The mos 36-item short-form health survey (sf-36). I. Conceptual framework and item selection. *Medical Care*, 30(6), 473–483. doi:10.1097/00005650-199206000-00002.
5. Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, 34(3), 220–233. doi:10.1097/00005650-199603000-00003.
6. Ware, J. E., Kosinski, M., & Keller, S. D. (1995). *How to score the sf-12 physical and mental health summary scales*. Boston: The Health Institute, New England Medical Center.
7. Stewart, A., Hays, R., & Ware, J. (Eds.). (1992). *Methods of validating the mos health measures*. London: Duke University Press.
8. Aaronson, N. K., Muller, M., Cohen, P. D., Essink-Bot, M. L., Fekkes, M., Sanderman, R., et al. (1998). Translation, validation, and norming of the Dutch language version of the sf-36 health survey in community and chronic disease populations. *Journal of Clinical Epidemiology*, 51(11), 1055–1068. doi:10.1016/S0895-4356(98)00097-3.
9. Williams, J. W., Jr., Pignone, M., Ramirez, G., & Perez Stellato, C. (2002). Identifying depression in primary care: A literature synthesis of case-finding instruments. *General Hospital Psychiatry*, 24(4), 225–237. doi:10.1016/S0163-8343(02)00195-0.
10. Composite international diagnostic interview (cidi). (1990). Geneva: World Health Organisation.
11. Smeets, R., & Dingemans, P. (1993). *Composite international diagnostic interview (cidi), version 1.1*. Amsterdam, The Netherlands: University of Amsterdam.
12. APA. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
13. Wittchen, H. U. (1994). Reliability and validity studies of the who—composite international diagnostic interview (cidi): A critical review. *Journal of Psychiatric Research*, 28(1), 57–84. doi:10.1016/0022-3956(94)90036-1.
14. Wittchen, H. U., Robins, L. N., Cottler, L. B., Sartorius, N., Burke, J. D., & Regier, D. (1991). Cross-cultural feasibility, reliability and sources of variance of the composite international diagnostic interview (cidi). The multicentre who/adamha field trials. *The British Journal of Psychiatry*, 159, 645–653, 658. doi:10.1192/bjp.159.5.645.
15. Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., et al. (1998). Cross-validation of item selection and scoring for the sf-12 health survey in nine countries: Results from the iqola project. International quality of life assessment. *Journal of Clinical Epidemiology*, 51(11), 1171–1178. doi:10.1016/S0895-4356(98)00109-7.
16. McHorney, C. A., Ware, J. E., Jr., Lu, J. F., & Sherbourne, C. D. (1994). The mos 36-item short-form health survey (sf-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, 32(1), 40–66. doi:10.1097/00005650-199401000-00004.
17. Ware, J. E., Jr. (1993). *Sf-36 health survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Centre.
18. Ware, J., Kosinski, M., & Keller, S. D. (1994). *Sf-36® physical and mental health summary scales: A user's manual*. Boston, MA: The Health Institute.
19. Lam, C. L., Tse, E. Y., & Gandek, B. (2005). Is the standard sf-12 health survey valid and equivalent for a Chinese population? *Quality of Life Research*, 14(2), 539–547. doi:10.1007/s11136-004-0704-3.
20. Ware, J. E., Jr., Gandek, B., Kosinski, M., Aaronson, N. K., Apolone, G., Brazier, J., et al. (1998). The equivalence of sf-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: Results from the iqola project International quality of life assessment. *Journal of Clinical Epidemiology*, 51(11), 1167–1170. doi:10.1016/S0895-4356(98)00108-5.
21. Jenkinson, C., & Layte, R. (1997). Development and testing of the uk sf-12 (short form health survey). *Journal of Health Services Research and Policy*, 2(1), 14–18.
22. Jenkinson, C., Layte, R., Jenkinson, D., Lawrence, K., Petersen, S., Paice, C., et al. (1997). A shorter form health survey: Can the sf-12 replicate results from the sf-36 in longitudinal studies? *Journal of Public Health Medicine*, 19(2), 179–186.
23. Kontodimopoulos, N., Pappa, E., Niakas, D., & Tountas, Y. (2007). Validity of sf-12 summary scores in a Greek general population. *Health and Quality of Life Outcomes*, 5, 55. doi:10.1186/1477-7525-5-55.

24. Jenkinson, C. (1999). Comparison of uk and us methods for weighting and scoring the sf-36 summary measures. *Journal of Public Health Medicine*, 21(4), 372–376. doi:[10.1093/pubmed/21.4.372](https://doi.org/10.1093/pubmed/21.4.372).
25. Farivar, S. S., Cunningham, W. E., & Hays, R. D. (2007). Correlated physical and mental health summary scores for the sf-36 and sf-12 health survey, v.I. *Health and Quality of Life Outcomes*, 5, 54. doi:[10.1186/1477-7525-5-54](https://doi.org/10.1186/1477-7525-5-54).
26. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
27. Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. doi:[10.1037/0033-2909.112.1.155](https://doi.org/10.1037/0033-2909.112.1.155).
28. Spindler, H., Kruuse, C., Denollet, J., & Pedersen, S. S. (2008). Positive and negative effect correlate differently with distress and health related quality of life in Danish cardiac patients: Cross-cultural validation of the global mood scale. *Psychosomatic Medicine* (revised version submitted).
29. Muller-Nordhorn, J., Roll, S., & Willich, S. N. (2004). Comparison of the short form (sf)-12 health status instrument with the sf-36 in patients with coronary heart disease. *Heart (British Cardiac Society)*, 90(5), 523–527. doi:[10.1136/hrt.2003.013995](https://doi.org/10.1136/hrt.2003.013995).
30. Schofield, M. J., & Mishra, G. (1998). Validity of the sf-12 compared with the sf-36 health survey in pilot studies of the Australian longitudinal study on women's health. *Journal of Health Psychology*, 3(2), 259–271. doi:[10.1177/135910539800300209](https://doi.org/10.1177/135910539800300209).
31. Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (Vol. 3). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
32. Jenkinson, C. (1998). The sf-36 physical and mental health summary measures: An example of how to interpret scores. *Journal of Health Services Research and Policy*, 3(2), 92–96.
33. Hann, M., & Reeves, D. (2008). The sf-36 scales are not accurately summarised by independent physical and mental component scores. *Quality of Life Research*, 17(3), 413–423. doi:[10.1007/s11136-008-9310-0](https://doi.org/10.1007/s11136-008-9310-0).
34. de Jonge, P., Denollet, J., van Melle, J. P., Kuyper, A., Honig, A., Schene, A. H., et al. (2007). Associations of type-d personality and depression with somatic health in myocardial infarction patients. *Journal of Psychosomatic Research*, 63(5), 477–482. doi:[10.1016/j.jpsychores.2007.06.002](https://doi.org/10.1016/j.jpsychores.2007.06.002).
35. Martens, E. J., Smith, O. R., & Denollet, J. (2007). Psychological symptom clusters, psychiatric comorbidity and poor self-reported health status following myocardial infarction. *Annals of Behavioral Medicine*, 34(1), 87–94. doi:[10.1007/BF02879924](https://doi.org/10.1007/BF02879924).
36. Crilley, J. G., & Farrer, M. (2001). Impact of first myocardial infarction on self-perceived health status. *Quarterly Journal of Medicine*, 94(1), 13–18. doi:[10.1093/qjmed/94.1.13](https://doi.org/10.1093/qjmed/94.1.13).
37. Augustovski, F. A., Lewin, G., Elorrio, E. G., & Rubinstein, A. (2008). The argentine-spanish sf-36 health survey was successfully validated for local outcome research. *Journal of Clinical Epidemiology*, 61(12), 1279–1284, e1276.
38. Hopman, W. M., Towheed, T., Anastassiades, T., Tenenhouse, A., Poliquin, S., Berger, C., et al. (2000). Canadian normative data for the sf-36 health survey. Canadian multicentre osteoporosis study research group. *Canadian Medical Association Journal*, 163(3), 265–271.