ELSEVIER

Journal of Clinical Epidemiology 53 (2000) 359–365

Journal of Clinical Epidemiology

Validity and reliability of the SF-36 Health Survey Questionnaire in patients with coronary artery disease

I. Failde*, I. Ramos

Escuela Universitaria de Ciencias de la Salud, Universidad de Cádiz, C/Duque de Nájera 18, 11002 Cádiz, Spain

Received 5 January 1999; received in revised form 20 August 1999; accepted 23 August 1999

Abstract

The validity and reliability of the SF-36 has been studied in 185 patients hospitalized with suspected ischemic cardiopathy, classified into four groups (AMI, unstable angina, nonischemic cardiologies, and absence of cardiologies). The validity of the construct has been analyzed by means of the association of the SF-36 with sociodemographic and clinical variables, and with diagnostic groups. The correlation of the subscales with GHQ-28 scores and the factorial structure have been studied. Internal consistency was measured by Cronbach's α and the item-internal consistency and item-discriminant validity. The validation result was as expected, although the scores were significantly lower in patients with unstable angina, only in the PF, VT, and GH subscales. The correlations with the GHQ-28 were high for MH and VT. The internal consistency was high (Cronbach's α 0.72–0.94). Factorial analysis identified eight factors, with the "anxiety" component of subscale MH remaining as an independent factor. These results suggest that the SF-36 is a useful scale for the differentiated clinical forms of ischemic cardiopathy, with the additional capability of reflecting the level of anxiety in these patients.

Keywords: Quality of Life; SF-36; Validity; Reliability, Coronary artery disease

1. Introduction

The increase observed in the survival of patients with ischemic cardiopathy (IC), together with the effect of the disease on the social, professional, and family life of those suffering it, have led researchers to consider that the traditional ways of measuring morbidity and mortality are not adequate for assessing the potential benefits of health care interventions. For this reason, there is common agreement on the need to use an indicator of subjective assessment of health, and of quality of life related to health (HRQL), as a complementary criterion for monitoring the results of medical interventions in these patients [1].

The term "quality of life" came into use during the 1970s as a multidimensional concept reflecting the overall subjective condition of the physical and mental welfare of the individual [2], which is a consequence not only of the disease but also of the family and social conditions forming the environment of the patient [3].

Generally speaking, in patients with acute myocardial infarction (AMI) and angina pectoris (AP), worse quality of life consequences have been observed than in other populations and differences have been detected as a function of the severity of the clinical condition [4–6]. The assessment of the quality of life of these patients has been approached by various authors both from disease-specific instruments [7,1] and from generic instruments such as the Nottingham Health Profile and the Sickness Impact Profile. Both types of instrument have advantages and disadvantages: specific questionnaires improve the sensitivity of measurement because they are specially designed to focus on the effects of a particular disease, whereas generic health measures allow comparison between different conditions and, because they quantify the patient's overall health, they may provide additional information [8–10].

The Short Form 36 (SF-36) is a generic instrument consisting of 36 items or questions grouped into eight health-related aspects of the patient's life. It has recently been adapted for Spanish-speaking patients [11] according to the protocol of the IQOLA project [12], and has been widely used in various countries with different population groups [9,10] and with both medical and psychiatric conditions [13]. Transversal studies have shown that the SF-36 is a valid and reliable instrument for detecting differences between groups defined by age, sex, socioeconomic status, and clinical condition [14,15], and its usefulness in patients with stable angina has been demonstrated [5,16]. However,

* Corresponding author. Tel.: +34-956212988; fax: +34-956224218.
E-mail address: inmaculada.failde@uca.es (I. Failde)

0895-4356/00/$ – see front matter
PII: S0895-4356(99)00175-4

no information exists on its validation and use in a population of patients with IC, in either the acute or chronic forms of the disease.

The General Health Questionnaire (GHQ) is a specific instrument considered to be a reference standard for the psychopathological screening of nonpsychiatric patients [17]. The version of this questionnaire with 28 items has been validated in Spanish by Lobo *et al*. [18] and assurance has been provided for its use as a means of detecting problems in cardiology patients [19] among whom are frequently found symptoms of conditions such as anxiety and depression that are rarely diagnosed by the cardiologist or else are inadequately attributed to physical disease.

The objective of this study was to ascertain the validity and reliability of the SF-36 in a Spanish population of patients admitted to hospital with suspected IC.

2. Methods

2.1. Subjects

The study was conducted in the University Hospital of Puerto Real (Cádiz, SW Spain), which covers a population of 218,813 inhabitants.

A total of 185 patients admitted into the Cardiology Service with suspected IC were included in the study. Patients were classified into four groups (G1 = AMI; G2 = unstable angina; G3 = nonischemic cardiologies; G4 = absence of cardiological disease), according to clinical criteria, electrocardiographic or biochemical data, and according to the medical report on release from hospital.

Patients were considered to have AMI (G1) if they met at least two of the following criteria: precordial pain of 20 minutes or more duration; CPK and CPK-MB above normal values in at least two serum samples; and/or the appearance of the Q wave in at least two ECG readings.

Patients were classified as having unstable angina if they suffered precordial pain similar to the first group, and showed changes in the ST segment of the ECG, without a high enzyme level. The third and fourth groups corresponded to patients in whom ischemic pathology (G3) and any cardiac pathology (G4), respectively, was discounted.

2.2. Instruments and analysis

The sociodemographic and clinical data were obtained from a structured questionnaire and from the clinical records of the patient. The classification devised by the Spanish Society for Epidemiology (SEE) was used as socioeconomic categories [20].

To assess the quality of life, the SF-36 questionnaire was used; from this, under each of the eight dimensions considered, item scores are coded, summed, and transformed to a scale ranging from 0 (worst health status) to 100 (best health status) [21]. The instrument used to record the mental health condition of the patient was the General Health Questionnaire-28 (GHQ-28), and a score of $\geqslant 6$ on this questionnaire was taken as the cutoff point; this provided a sensitivity of 76.9% and a specificity of 90.2% [22].

In order to validate the construct, the hypothesis adopted was that the HRQL in the population studied would be associated inversely with the patient's age, with lower socioeconomic level, with worse mental health condition (GHQ <6), with the chronicity of the IC (unstable angina vs. AMI), and the presence of other pathologies or of a number of cardiovascular risk factors. Moreover, based on other studies [15], it was considered that the HRQL would show relatively worse results for female patients.

For the analysis of the construct, the mean and SD were calculated for each of the subscales of the SF-36, using as constructs the previously listed variables. The Student *t*-test and one-way ANOVA ($P < 0.05$) were used for comparison of groups, while the Tukey-Kramer test was used for the comparison of groups 2 by 2. The percentage of patients achieving the lowest (floor) and the highest (ceiling) possible score were also computed for the whole group of patients.

Principal component analysis followed by Varimax rotation was performed to determine the structure of the SF-36 in our population. A factor was considered relevant only if its eigenvalue exceeded 1 [23].

Concurrent validity was tested by the Pearson's correlation ($\pm$CI 95%), considering that the mental health as measured by the GHQ-28 should be more closely relate to the similar dimensions of the SF-36 (mental health, emotional role, vitality, and social functioning) than to the other dimensions (PF, RF, and BP [13]).

Two techniques were used to assess the internal consistency of the SF-36: item-scale correlations (item-internal consistency and item-discriminant validity) [24] and Cronbach's α. Item-internal consistency, which assesses the extent to which an item is related to the remainder of its scale with the item omitted, should exceed 0.4 [25]. Furthermore, items should be more closely related to their own scale than to the other scales (item-discriminant validity). A scaling success rate was computed as a summary of the item-discriminant validity of each scale. A scaling success was counted for each item whenever it correlated more highly with its hypothesized scale than with all the other scales of the questionnaire [26]. Cronbach's α measures the overall correlations between items within a scale. Reliability is considered acceptable for group comparisons when α exceeds 0.7 [27].

The computer programs Epinfo v.6.0 and SPSS v.6 for Windows were used for the tabulation and statistical analysis of the data.

3. Results

3.1. Characteristics of the population

A total of 185 patients was studied, 71% of whom were males. Of the total, 68.1% were aged over 56 years (mean

Table 1
Descriptive statistics of score distributions for SF-36 scales: combined sample

	PF	RP	BP	GH	VT	SF	RE	MH
Mean	77.1	66.6	67.3	57.2	62.7	79.2	60.5	65.7
SD	25.0	42.1	31.5	20.0	24.2	26.9	46.2	24.4
Minimum	5	0	0	5	5	0	0	4
Maximum	100	100	100	100	100	100	100	100
% Floor	0	22	0.5	0	0	0.5	33	0
% Ceiling	28	55	39	3	8	49	65	6

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

60.2 years, SD = 10.4) and 55% were in the socioeconomic categories of semiqualified or unqualified manual workers (Groups IV and V of the SEE classification). Clinically, the diagnostic group least represented in the sample was that of patients with nonischemic cardiopathy (10.8%). It was also observed that 63.2% of the patients studied presented one or more pathologies associated with the base disease, and that 74.6% of those had at least one cardiovascular risk factor. In addition, 49.2% of patients had GHQ-28 scores of ⩾6 points (means 6.25, SD = 5.2). Differences between the diagnostic groups were detected only in the variable of sex, where more males were observed in the groups with AMI and unstable angina. Also, the factor of personal history of the disease was more frequent in the diagnostic group with unstable angina (60%).

The average values for the whole population of the various dimensions of the SF-36 were lowest for the GH and highest for the SF subscales (Table 1).

The proportion of patients with best possible scores or "ceiling effect" was high for the RE, RP, SF, and BP subscales. The proportion of worst possible scores or "floor effect" was also high for the RE and RP subscales (Table 1).

3.2. Construct validity

Tables 2 and 3 show the means scores (SD) of the subscales of the SF-36 and the different variables used as constructs. It can be seen in Table 2 that females score significantly lower in all the subscales except BP. However, the older patients only showed significantly lower scores (P = 0.01) in subscale PF. In other dimensions such as RE and SF, younger patients showed lower scores but these were not statistically significant.

The worst scores in all the subscales, excepting BP, were shown by housewives, who were the main constituent of the "unclassifiable" socioeconomic group, followed by the more disadvantaged groups (Table 2).

In Table 3 it can be seen that the patients in the group diagnosed with unstable angina (G2) scored significantly worse in the PF and GH subscales. However, the group of patients without cardiopathy (G4) scored worse in the VT subscale.

In the comparison by groups, significant differences were observed only between patients with AMI (G1) and unstable angina (G2) in the PF, VT, and GH subscales, although the scores were lower in all the subscales for the unstable angina group (G2).

When the presence of other pathologies and of cardiovascular risk factors was studied, it was observed that, in the PF and GH subscales, the scores were significantly worse the more comorbidity or risk factors that were present, with the RP subscale included with these, to the extent that more risk factors were present (Table 3).

Patients with GHQ-28 scores of ⩾6 presented worse results in all the subscales of the SF-36 (Table 3).

Table 2
Mean scores (SD) of the SF-36 and sociodemographic variables of the population studied

	n (%)	PF	SF	RP	RE	MH	BP	VT	GH
Age									
<46	18 (9.7)	87.5 (15.1)	68.1 (34.3)	75.0 (37.3)	46.2 (47.3)	62.4 (29.3)	64.1 (34.7)	61.1 (34.3)	61.5 (26.1)
46-55	41 (22.2)	81.5 (25.2)	79.8 (25.2)	75.0 (39.5)	64.2 (47.7)	62.5 (28.8)	63.2 (32.7)	60.4 (23.2)	52.1 (21.6)
56-65	62 (33.5)	79.1 (24.4)	81.6 (26.3)	65.3 (42.3)	59.1 (46.1)	69.1 (21.3)	72.8 (29.0)	66.0 (24.1)	57.5 (18.4)
>65	64 (34.6)	69.3 (25.9)	79.4 (26.2)	60.1 (44.2)	63.5 (45.8)	65.3 (22.6)	65.2 (32.1)	61.2 (21.8)	59.1 (18.3)
Significance		0.01	0.31	0.27	0.50	0.53	0.39	0.61	0.26
Sex									
Males	133 (71.9)	80.6 (23.2)	82.3 (25.2)	72.7 (39.1)	69.6 (42.5)	69.7 (23.2)	68.2 (31.4)	67.7 (21.8)	59.3 (20.5)
Females	52 (28.1)	68.1 (27.3)	71.1 (29.6)	50.9 (45.6)	37.1 (47.4)	55.3 (24.3)	64.9 (31.9)	49.8 (25.5)	52.2 (17.8)
Significance		0.005	0.01	0.003	0.000	0.000	0.53	0.000	0.02
Socioeconomic level[a]									
I & II	13 (7.1)	96.9 (5.2)	97.1 (5.4)	88.4 (19.4)	87.1 (32.0)	80.6 (22.7)	81.4 (27.1)	75.3 (21.9)	76.7 (15.1)
III	31 (16.9)	78.3 (23.7)	78.6 (29.2)	66.9 (43.4)	62.3 (46.9)	63.7 (24.4)	66.8 (30.0)	62.2 (25.6)	52.9 (22.1)
IV & V	97 (53.0)	77.9 (24.1)	79.8 (25.9)	70.6 (39.8)	66.6 (43.0)	68.0 (25.1)	66.8 (32.1)	65.5 (23.3)	58.1 (19.8)
Not classifiable	42 (23.0)	68.5 (27.7)	71.7 (29.7)	48.8 (46.5)	34.9 (47.6)	58.0 (22.7)	64.5 (31.9)	53.4 (23.7)	52.9 (16.9)
Significance		0.003	0.02	0.007	0.000	0.01	0.39	0.01	0.001

[a]I = senior managers, professionals; II = managers, technically qualified, and trades people; III = supervisory and administrative; IV = skilled manual workers; V = unskilled manual workers; Not classifiable = housewives and nonmembers of workforce.

Table 3
Mean scores (SD) of the SF-36 and clinical variables of the population studied

	n (%)	PF	SF	RP	RE	MH	BP	VT	GH
Diagnostic groups									
G1	62 (33.5)	83.1 (23.4)[a]	83.1 (22.9)	76.2 (37.7)	63.4 (46.2)	70.6 (23.2)	74.5 (30.2)	70.3 (21.4)[a]	64.6 (18.1)[a]
G2	70 (37.8)	69.9 (26.7)[a]	77.3 (28.2)	61.4 (43.5)	54.2 (46.8)	62.0 (24.8)	63.1 (30.9)	59.9 (24.0)[a]	51.7 (18.7)[a]
G3	20 (10.8)	79.7 (23.6)	83.7 (26.3)	51.2 (45.5)	56.6 (47.2)	69.8 (21.9)	64.1 (35.2)	64.5 (24.7)	54.3 (24.3)
G4	33 (17.9)	79.3 (21.9)	73.1 (30.9)	68.9 (41.9)	70.7 (43.9)	61.8 (25.8)	64.3 (31.8)	53.0 (26.0)	57.3 (19.8)
Significance		0.01	0.27	0.06	0.35	0.13	0.17	0.005	0.002
Comorbidity									
No pathology	63 (36.8)	83.1 (23.9)	78.3 (27.5)	72.4 (40.1)	66.2 (45.5)	67.5 (27.4)	73.9 (30.1)	65.2 (25.2)	62.7 (19.4)
1 pathology	43 (23.2)	79.8 (26.4)	78.7 (25.3)	65.1 (44.3)	62.0 (45.2)	63.5 (21.6)	61.4 (31.2)	60.3 (24.3)	55.8 (16.7)
⩾2 pathologies	74 (40.0)	69.9 (23.6)	80.2 (27.6)	62.1 (42.4)	54.5 (47.3)	65.3 (23.0)	64.5 (32.3)	61.8 (23.4)	53.1 (21.4)
Significance		0.004[b]	0.91	0.33[b]	0.31	0.69[b]	0.07[b]	0.53[b]	0.01[b]
Risk Factors									
No RF	47 (25.4)	87.4 (17.7)	81.6 (23.2)	82.4 (32.9)	64.5 (44.7)	73.0 (23.3)	75.1 (28.7)	69.2 (20.8)	65.4 (21.3)
1 RF	75 (40.5)	76.4 (26.2)	80.6 (27.8)	63.0 (43.9)	62.2 (45.9)	64.8 (23.3)	67.1 (32.2)	62.4 (26.0)	58.0 (19.1)
2 RF	46 (24.9)	73.4 (26.6)	76.9 (29.3)	63.0 (43.3)	59.4 (48.1)	60.7 (27.2)	64.1 (32.6)	59.7 (24.7)	51.1 (18.1)
⩾3 RF	17 (9.2)	62.8 (21.4)	71.1 (31.8)	45.3 (42.0)	41.6 (46.3)	63.7 (20.6)	57.1 (29.2)	54.6 (21.0)	48.8 (17.4)
Significance		0.002	0.50	0.008	0.37	0.09	0.16	0.11	0.001
GHQ-28 score									
<6	94 (50.8)	87.4 (17.7)	90.0 (19.8)	82.4 (33.7)	77.3 (40.1)	77.7 (17.8)	76.5 (28.7)	74.8 (20.3)	67.1 (15.6)
⩾6	91 (49.2)	66.3 (26.7)	67.9 (28.8)	50.2 (43.6)	43.2 (45.9)	53.2 (24.0)	57.7 (31.6)	50.1 (21.4)	47.0 (19.0)
Significance		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

[a]Tukey-Kramer test: $p < 0.01$.
[b]*F* test for linear trend: $p < 0.05$.

The factorial analysis identified eight factors with eigenvalues between 10.6 and 1.1 (Table 4). The items of the subscale MH were shared between factor 2 and factor 6; whereas factor 2 consisted of the items of subscales VT and SF as well as those of subscale MH that explore aspects such as sadness/happiness, in factor 6 the items of MH that explore states of anxiety/calm remained independent.

The five items of subscale GH were distributed between factors 5 and 8, with those items representing a subjective evaluation of health being differentiated from those assessing the degree of illness.

Lastly, in the correlations made between the GHQ-28 and the subscales of the SF-36 (Table 5), it was observed that they were higher for the MH and VT subscales, and lower for BP, RP, and PF. However, these coefficients were only statistically significant for BP with respect to MH and VT.

3.3. Internal consistency

The correlation coefficients between items and the remainder of their own scale ranged from 0.53 to 0.95, and they were all higher than with other scales (Table 6), providing further evidence of internal consistency. Perfect scaling success rates were obtained across all the SF-36 scales (Table 6).

For all eight subscales, internal consistency measured by Cronbach's α exceeded 0.7 (Table 6).

4. Discussion

The SF-36 is considered to be a valid, reliable, concise generic measure of state of health that is potentially useful for application to groups of patients. However, its psychometric properties in patients with ischemic cardiopathy has not previously been studied properly. The objective of this study was to assess the validity and reliability of the SF-36 in a population of patients hospitalized with suspected ischemic cardiopathy.

The results obtained indicate that the SF-36 is a useful measure for assessing the quality of life in these patients, although its usefulness could be limited for some of the subscales comprising the instrument. On this point, one should note the large SDs found in the subscales RE and RP, which have also been observed by other authors in other populations [9]. This would be explained by the large number of subjects with high percentages in the extreme values of the scale (the ceiling and floor effect), a fact already stated by authors such as McHorney *et al*. [27], who consider that these two subscales are the most "coarse" of all the eight, and who discuss the possibility of including more response categories for the items comprising these two subscales, in order to be able to establish a graduation in role disability, rather than the mere presence or absence of disability.

With respect to the construct validity, the scale behaves as expected, and as reflected by other studies, in respect of the sociodemographic variables [10,28] with the exception of age, where a significant descending gradient is only observed in subscale PF. Authors such as Hemingway *et al*. [15], in the Whitehall II study conducted in an large sample of ill and healthy subjects, find results similar to ours, and even find a significant increase with age in the subscales MH, RE, VT, and SF. The authors interpret their results with caution, indicating the possibility of a cohort effect. At

Table 4
Factorial Analysis of the SF-36 in the population studied

	Factor coefficients of individual items after rotation							
	F1	F2	F3	F4	F5	F6	F7	F8
Physical Functioning (PF)								
PF1	**0.59**	0.12	0.16	−0.11	0.40	−0.14	0.18	−0.06
PF2	**0.81**	0.08	0.03	0.03	0.17	−0.04	0.09	−0.04
PF3	**0.77**	0.15	0.08	0.07	0.18	0.02	0.02	0.02
PF4	**0.74**	0.21	0.15	0.03	0.30	0.01	0.04	−0.12
PF5	**0.78**	0.06	0.10	0.02	0.07	0.10	−0.02	−0.2
PF6	**0.69**	0.21	0.08	−0.07	0.23	−0.04	0.09	0.01
PF7	**0.78**	0.12	0.16	0.13	0.03	0.10	0.03	0.13
PF8	**0.79**	0.14	0.09	0.09	−0.20	0.02	−0.08	0.29
PF9	**0.80**	0.06	−0.00	0.06	−0.20	0.01	−0.04	0.27
PF10	**0.53**	−0.06	0.20	0.09	0.03	0.25	0.16	0.25
Mental health (MH)								
MH1	0.01	0.26	0.14	0.20	−0.03	**0.78**	−0.05	0.10
MH2	0.16	**0.72**	0.02	0.20	0.13	0.13	0.07	0.00
MH3	0.04	0.27	0.03	0.09	0.06	**0.80**	0.09	0.06
MH4	0.13	**0.75**	0.06	0.14	0.08	0.16	0.07	0.19
MH5	0.11	**0.74**	0.17	0.13	0.11	0.07	0.08	0.14
Social functioning (SF)								
SF1	0.03	**0.61**	0.36	0.20	−0.03	0.03	0.09	0.09
SF2	0.11	**0.69**	0.13	0.32	−0.12	0.14	0.14	0.10
Vitality (VT)								
VT1	0.16	**0.63**	0.14	0.04	0.29	0.09	0.07	0.07
VT2	0.20	**0.45**	0.27	0.02	0.38	0.13	−0.00	−0.25
VT3	0.25	**0.44**	0.26	0.06	0.27	0.33	0.02	−0.15
VT4	0.29	**0.46**	0.10	0.03	0.31	0.34	0.07	−0.02
Role—physical (RP)								
RP1	0.15	0.13	**0.85**	0.11	0.08	−0.01	0.03	−0.02
RP2	0.13	0.15	**0.87**	0.10	0.10	−0.01	0.05	0.04
RP3	0.14	0.19	**0.84**	0.10	0.11	0.13	0.07	0.06
RP4	0.17	0.18	**0.78**	0.14	0.07	0.14	0.11	0.08
Role—emotional (RE)								
RE1	0.09	0.29	0.19	**0.85**	0.08	0.15	0.06	−0.10
RE2	0.06	0.26	0.17	**0.87**	0.08	0.10	0.09	0.06
RE3	0.06	0.21	0.13	**0.88**	0.11	0.08	−0.03	−0.07
General health (GH)								
GH1	0.24	0.13	0.15	0.13	**0.59**	−0.03	0.22	0.31
GH2	0.19	0.31	0.10	0.06	0.18	0.03	0.09	**0.67**
GH3	0.07	0.24	0.07	0.08	**0.74**	0.02	0.10	0.11
GH4	0.04	0.06	0.04	−0.11	0.30	0.12	−0.05	**0.59**
GH5	0.16	0.05	0.10	0.15	**0.68**	0.08	0.11	0.30
Bodily pain (BP)								
BP1	0.01	0.15		0.06	0.10	0.07	**0.89**	0.09
BP2	0.16	0.22		0.03	0.23	−0.01	**0.82**	−0.08
Eigenvalue before rotation	10.6	3.9	2.2	2.1	1.5	1.4	1.2	1.1

the same time, they refer to prospective studies demonstrating the tendency of successive generations to report their health more adversely [29].

The subscales that concern aspects that are more physical than mental and emotional were more sensitive in detecting differences in the presence or absence of comorbidity or of risk factors. McHorney *et al*. [13], in this respect, report that subscale PF is more sensitive for distinguishing groups that differ in the severity of chronic medical condition. Similarly, Brazier *et al*. [10] do not find significant differences in the MH and RE subscales according to the presence or absence of chronic problems.

In the evaluation between groups, we observe that only in the subscales PF, VT, and GH were statistically significant differences detected between the groups of patients with AMI and unstable angina, possibly due to the more frequent presence of personal history of the disease among the latter group. The absence of differences in other subscales such as MH and RE could be explained by the presence of associated organic pathology and particularly of psychiatric disorders in the two groups. Spertus *et al*. [1] believe that the presence of other illnesses could give rise to a change in the quality of life, not necessarily in the same direction as that produced by the IC, and that

Table 5
Correlations of the dimensions of the SF-36 with the total score of the population in the GHQ-28

SF-36 subscales/dimensions	Correlation with GHQ-28 ± CI 95%
Physical functioning (PF)	−0.46* (−0.34; −0.57)
Social functioning (SF)	−0.53* (−0.42; −0.63)
Role limitation attributed to physical problems (RP)	−0.44* (−0.31; −0.55)
Role limitation attributed to emotional problems (RE)	−0.51* (−0.39; −0.61)
Mental health (MH)	−0.61* (−0.51; −0.69)
Vitality (VT)	−0.61* (−0.51; −0.69)
Bodily pain (BP)	−0.35* (−0.22; −0.47)
General health perception (GH)	−0.58* (−0.47; −0.68)

Correlation coefficients are negative because two scales run in opposite directions. CI 95% = $1/2 \ln (1 + r/1 - r) \pm z_{\alpha/2} (1/\sqrt{n-3})$

*$p = 0.000$.

the influence of such other illnesses on the general health of patients could limit the sensitivity of the SF-36 in detecting differences. However, and supporting the capacity of the SF-36 to identify changes in the mental health condition, a moderately high correlation was found between the GHQ-28 and the MH subscale of the SF-36, as well as the capacity of the scale to discriminate between subjects with GHQ-28 scores of <6 and ≥ 6. In this respect, our results are supported by those presented by other authors [30,31] and could be explained by the effect of mental disorders on social life, physical limitation, and emotional well-being.

Furthermore, our findings supported the developer's claim of internal consistency for the SF-36 questionnaire [27], because all the coefficients were at values above those recommended, and the results of both item-internal consistency and item-discriminant validity were similar to those found by McHorney *et al*. [27] in different groups of patients.

Regarding the factorial structure of the scale, our results are close to those found by the original authors in four of the eight SF-36 scales; however, the grouping of the items in the MH and GH subscales in our population is different from that proposed. The SF-36 in our patients shows a marked difference between the states of anxiety/calm and sadness/happiness, these latter states being related with vitality and social functioning. This is somewhat similar to that observed by Garrat *et al*. [24] between the subscales VT and MH, and could be explained by the tendency of these patients to relate sadness to lack of vitality (mean $r = 0.49$) and lack of vitality to difficulties in social relationships (mean $r = 0.41$). Moreover, the greater weight of the component of anxiety in the overall state of mental health of this population (data not presented) could explain the independence observed in the items that explore this factor; in addition, this observation supports the capacity of the scale for identifying this state in the population.

On the basis of all the foregoing, and as the conclusion of this article, it can be stated that the SF-36 is a useful scale for evaluating the quality of life in patients with different manifestations of IC. Further, the SF-36 is capable of reflecting special characteristics of this population, such as the level of anxiety affecting them. However, there are limitations in certain subscales, such as RE, already expressed by other authors with respect to other populations, particularly when the objective is to establish differences between diagnostic groups.

Acknowledgment

This study received financial support from the Spanish Ministry of Health (Fondo de Investigación Sanitaria 99/1036).

References

[1] Spertus JA, Winder JA, Dewhurst TA, Deyo RA, Fihn SD. Monitoring the quality of life in patients with coronary artery disease. Am J Cardiol 1994;74:1240–4.
[2] Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Boston, MA: Health Institute, New England Medical Center, 1993.

Table 6
Internal consistency and homogeneity of the SF-36 scale of the population studied

Dimensions	Number of items	Range of item correlations: Item-internal consistency[a]	Range of item correlations: Item-discriminant validity[b]	Scaling success rate (%)	Cronbach's α
PF	10	0.55—0.84	0.03—0.45	100	0.92
SF	2	0.88—0.89	0.21—0.51	100	0.72
RP	4	0.86—0.91	0.18—0.45	100	0.91
RE	3	0.93—0.95	0.16—0.52	100	0.94
MH	5	0.73—0.80	0.05—0.61	100	0.80
VT	4	0.76—0.80	0.18—0.44	100	0.78
BP	2	0.92—0.93	0.13—0.35	100	0.82
GH	5	0.53—0.78	0.07—0.43	100	0.74

[a] Correlations between items and it own scale corrected for overlap.
[b] Correlations between items and others scale.

[3] McSweeny AJ, Creer TL. Health-related quality of life assessment in medical care. Disease-a-Month 1995;XLI:3–71.
[4] Rawles J, Light J, Watt M. Quality of life in the first 100 days after suspected acute myocardial infarction-a suitable trial endpoint? J Epidemiol Community Health 1992;46:612–6.
[5] Charlier L, Dutrannois J, Kaufman L. The SF-36 questionnaire: a convenient way to assess quality of life in angina pectoris patients. Acta Cardiol 1977;3:247–60.
[6] Marquis P, Fayol C, Joire JE. Clinical validation of quality of life questionnaire in angina pectoris patients. Eur Heart J 1995;16:1554–60.
[7] Wilson A, Wiklund I, Lathi T, Wahl M. A summary index for the assessment of quality of life in angina pectoris. J Clin Epidemiol 1991; 44:981–8.
[8] Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369–78.
[9] Schlenk EA, Erlen JA, Dunbar-Jacob J, McDowell J, Engberg S, Sereika SM, Rohay JM, Bernier MJ. Health-related quality of life in chronic disorders: a comparison across studies using the MOS SF-36. Qual Life Res 1998;7:57–65.
[10] Brazier JE, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992;305: 160–4.
[11] Alonso J, Prieto L, Antó JM. La versión española del SF-36 Health Survey (cuestionario de salud SF-36): un instrumento para medida de los resultados clínicos. Med Clin (Barc) 1995;104:771–6.
[12] Bullinger M, Alonso J, Apolone G, Leplége A, Sullivan M, Wood-Dauphine S. *et al*. Translating health status questionnaires and evaluating their quality: The IQOLA Project approach. J Clin Epidemiol 1998;51:913–23.
[13] McHorney CA, Ware JE, Raczek AE. The MOS 36 item Short-Form Health Survey (SF-36): II. Psychometric and clinical test of validity in measuring physical and mental health constructs. Med Care 1993; 31:247–63.
[14] Hemingway H, Stafford M, Stansfeld S, Shipley M, Marmot M. Is the SF-36 a valid measure of change in population health? Results from Whitehall II study. BMJ 1997;315:1273–9.
[15] Hemingway H, Bchir MB, Nicholson A, Stafford M, Roberts R, Marmot M. The impact of socioeconomic status on health functioning as assessed by the SF-36 questionnaire: the Whitehall II study. Am J Public Health 1997;87:1484–90.
[16] Permanyer-Miralda G, Alonso J, Antó JM, Alijarde-Guimerá M, Soler-Soler J. Comparison of perceived health status and conventional functional evaluation in patients with stable coronary artery disease. J Clin Epidemiol 1991;44:779–86.
[17] Goldberg DP, Rickels K, Downing R, Hesbacher P. A comparison of two psychiatric screening tests. Br J Psychol 1976;129:61–7.
[18] Lobo A, Perez MJ, Artal J. Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. Psychol Med 1986;16:135–40.
[19] Vazquez-Barquero JL, Ochoteco A, Padierna JA, Antón JA, Arenal A, Peña C. El cuestionario de salud general (GHQ) como instrumento de "screening" de patología psiquiátrica en enfermos cardiológicos. Actas Luso-Esp Neurol Psiquiatr 1982;10:21–32.
[20] Grupo de trabajo de la Sociedad Española de Epidemiología. La medición de la clase social en ciencias de la salud. Barcelona: SG editores, 1995.
[21] Jenkinson C, Layte R, Wright L, Coulter A. The U.K. SF-36: an analysis and interpretation manual. A guide to health status measurement with particular reference to the Short Form 36 Health Survey. Health Services Research Unit, Departament of Public Health and Primary Care, University of Oxford, 1993.
[22] Badía J, Salamero M, Alonso J, Ollé A. La medida de la salud. Guía de escalas de medición en español. Barcelona: PPU, 1996.
[23] Joliffe IT. Principal Components Analysis. New York: Springer-Verlag, 1996.
[24] Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF-36 health survey questionnaire: an outcome measure suitable for routine use within the NHS?. BMJ 1993;306:1440–4.
[25] Streiner DL, Norman GR. Health Measurement Scales: a Practical Guide to their Development and Use. Oxford: Oxford University Press, 1992.
[26] Prieto L, Alonso J, Ferrer M. Are results of the SF-36 Health Survey and the Nottingham Health Profile similar?: A comparison in COPD patients. J Clin Epidemiol 1997;50:463–73.
[27] McHorney CA, Ware JE, Rachel Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Test of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40–66.
[28] Jenkinson C, Coulter A, Wright L. Short form 36 (SF 36) health survey questionnaire: normative data for adults of working age. BMJ 1993;306:1437–40.
[29] Barsky AJ. The paradox of health. N Engl J Med 1988;318:414–8.
[30] Ayuso-Mateos JL, Lasa L, Vazquez-Barquero JL, Oviedo A, Diez-Manrique JF. Measuring health status in psychiatric community surveys: internal and external validity of Spanish version of the SF-36. Acta Psychiatr Scand 1999;99:26–32.
[31] Spitzer RL, Kroenke K, Linzer M, Hahn SR, Williams JB, deGruy FV 3rd *et al*. Health-related quality of life in primary care patient with mental disorders. Results from the PRIME-MD 1000 study. JAMA 1995;1975:1511–7.