Short Form 36-item Health Survey (SF-36) in patients with physical disabilities: validity and reliability

Nassim Ahmadi[1], Mojgan Rajati[2], Maryam Khazaei Pool[3,4], Leili Salehi[5,6], Mohammad Ebrahimi[7], Fatemeh Rajati[8]*

1. Department of Speech and Language Pathology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

2. Department of Gynecology and Obstetrics, School of Medicine, Kermanshah University of Medical Sciences, Kermanshah, Iran.

3. Department of Public Health, School of Health, Mazandaran University of Medical Sciences, Sari, Iran.

4. Health Sciences Research Center, Addiction Research Institutes, Mazandaran University of Medical Sciences, Sari, Iran.

5. Department of Health Education and Promotion & Research center for Health, Safety and Environment, Alborz University of Medical Sciences, Karaj, Iran

6. Department of Health Education & Promotion, School of Public Health, Alborz University of Medical Sciences, Karaj, Iran

7. Student Research Committee, Kermanshah University of Medical Sciences, Kermanshah, Iran

8. Research Center for Environmental Determinants of Health, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

*Corresponding Author:
Research Center for Environmental Determinants of Health, Institute of Health, Kermanshah Medical Sciences University, Kermanshah, Iran.
Tel: +9838362007

Abstract

Aim: The aim of this study was to assess the performance of the SF-36 health survey (SF-36) in a sample of subjects with physical disabilities.

Material and Methods: 305 patients recruited using the convenient sampling method from September 2019 to March 2020 in Kermanshah, Iran. Another similar 300 patients were selected to assess the confirmatory factor analysis (CFA).

Results: The Cronbach's α ranged from 0.70 to 0.93, and intra-class correlation coefficients from 0.71 to 0.88; and with the no ceiling and floor effect for two main subscales. Convergent validity was supported by moderate to good correlation between SF-36 subscales and Moorong self-efficacy subscales (r= 0.25- 0.53). The SF-36 divergently validated with HADS total score and subscales (r= -0.24- -0.65), concurrently validated with its subscales (r=0.49- 0.88), and physically discriminated between persons with different level of disability (t-test: p<0.001). Factorial analysis identified seven factors, confirmed with second-order in another 300 samples (chi-square ($\chi 2$/df) = 2.61(p < .001); RMSEA = 0.07 (90% CI = 0.07–0.08); AGFI=0.75; GFI = 0.78; CFI= 0.85; and NFI = 0.78).

Conclusions: the SF-36 is a reliable and valid tool in physical disables. However, SF-36 shows insufficient eight-factor validity. Future studies should focus on evaluating other psychometric properties, such as sensitivity to change in subjects with physical disabilities.

Keywords: Disabled Persons, Health Status, Iran, Psychometrics, Quality of Life, Reliability

Introduction

Quality of life is commonly affected in persons with physical disabilities. Global Burdon of Disease (GBD) estimates 54·2 million (47·4–61·6) road injuries in 2017, which is one of the cause of the disabilities[1]. Individuals with physical disabilities experience various barriers due to restricted activity and their lives are affected with poor health outcomes[2]. It is reported that people with musculoskeletal dysfunction had the worst perception of their quality of life [3]. Generic health questionnaires that are developed to assess quality of life in general population can be used across various diseases, so assessing their psychometric properties in various diseases is common and valuable[4]. Among the several generic measures, the General Health Survey Short-Form-36 (SF-36) is a well-documented questionnaire for evaluation of health-related quality of life (HRQoL) in both healthy people and patients with different physical disability [5-8]. A Persian version of SF-36 validated by Montazeri et al. (2005) and shown to be valid and reliable[9]. As a well-tailored quality of life questionnaire can discriminate and compare between patients and healthy population [10,11], predict other comorbidities in persons with physical disabilities over time[12], and evaluate discomfort in this group[13,14], evaluating the validity and reliability of the SF-36 in subjects with physical disabilities is necessary. Although the SF-36 has been validated in different population, its validity has not yet been examined among Patients with physical disability. In current study we investigated validity and reliability of Persian version of Sf-36 in Iranian people with physical disabilities.

Materials and Methods

Participants and design

This validation study was conducted at four physiotherapy clinics in Kermanshah, Iran, from September 2019 to March 2020. A total of 305 subjects with physically disabled were

recruited, using the convenient sampling method from four rehabilitation centers. Also, 120 physically normal subjects from their families entered study for the purpose of validity analysis. Another 300 subjects with physical disabilities were recruited for confirmatory factor analysis from the same rehabilitation centers using convenient sampling method.

Inclusion criteria for current study were as follows: (a) subjects aged more than 18 years; and (b) people with any loco-motor disability, amputation, paralysis, deformity or dysfunction of limbs. The following exclusions were made: (a) insufficient cognitive abilities, (b) sensory motor, mental, emotional problems, and chronic condition affecting quality of life including diabetes, cardiovascular disease and cancer.

After written informed consent, the first author explained the purpose of the study and the questionnaires were administered to each subject and were filled out. A demographic questionnaire, the 36-Item Short-Form Health Survey (Sf-36), Hospital Anxiety and Depression Scale (HADS) were distributed to subjects. Participants were asked to complete the study questionnaires on the same day and were allowed to stop between completing questionnaires if they needed rest. individuals with physical disabilities completed the questionnaire independently. Help was available in answering to questions by the first author. Each questionnaire received a code and did not show any identifying information about the subjects. Participation in the study was totally voluntary and all data were kept confidential. The study was approved by the Ethics Committee of the Kermanshah University of Medical Sciences, with the ethic code of IR.KUMS.REC.1398.899.

Questionnaires

- *Demographic questionnaire***.** The demographic questionnaire comprised socio-demographic information including age, sex, educational level, job, aids, comorbidity, marital status, duration of disability, severity of disability, physical activity level,

- *Persian version of 36-Item Short-Form Health Survey (SF-36):* The 36-Item Short Form Health Survey (SF-36) questionnaire was developed by Ware and Sherbourne[15] . The SF-36 questionnaire is mainly used to evaluate the quality of life in the general population. This instrument measures eight health related domains: physical functioning (PF-), role physical (RP-), bodily pain (BP), general health (GH-), vitality (VT-), social functioning (SF-), role emotional (RE-), and mental health (MH-) [15]. Among the eight subscales, four subscales including PF, RP, BP and GH contribute to physical component summery (PCS) and the other four subscales including VT, SF, RE and MH are related to mental component summery (MCS). These subscales evaluate the associated quality of life within a range of 0-100 points with 0 point indicating poor and 100 points good quality of life [15,9]. Translation and validation of the Persian version of SF-36 has been performed by Montazeri et al (2005) and showed to be a valid and reliable measure, the Cronbach's α coefficients ranging from 0.77 to 0.90 with the exception of the vitality subscale (α= 0.65) [9]. Second question of SF-36 evaluates health trend during last year and is not counted in scoring,

- *Hospital Anxiety and Depression Scale (HADS):* HADS is a self-assessment screening scale which was developed by Zigmond and snaith (1983) to discover states of depression, anxiety and emotional distress among various patients [16]. HADS scale has two factors, the anxiety and depression sub-scales, both internally consistent, with Cronbach's α coefficient 0.80 and 0.76, respectively. The possible scores for each subscale range from 0 to 21, categorized as follows: normal 0–7, mild 8–10, moderate 11–14, and severe 15–21 [16].The reliability of HADs scale has been established via Crobanch α as 0.80–0.93 in patients with physical disorder[17]. The Persian version of HADS is validated previously with Cronbach's alpha as 0.78 for anxiety and 0.86 for depression subscale [18].

- *Moorong Self-Efficacy Scale (MSES):* The MSES was designed to assess the self-efficacy in activities of daily living in patients with spinal cords injury[19]. Daily activities and social functioning are two subscales of this questionnaire. Scoring of this questionnaire is based on an a 7-point scale, ranging from 1 (very uncertain) to 7 (very certain). Possible scores of this scale range from 16 to 112 [19]. Persian version of the Moorong Self-Efficacy Scale among subjects with physical disability was validated with Cronbach alpha coefficients of 0.94 which shows an excellent internal consistency [20].

Data Analysis

- *Statistical Analysis*

A descriptive statistic obtained for all variables and described as analyses were carried out using SPSS 22 (SPSS Inc., Chicago, IL, USA). Data were described as mean ± standard deviation (M±SD) for the quantitative variables and frequency presents the categorical variables. The Kolmogorov-Smirnov test was used for evaluation of normal distribution for quantitative variables. Statistically significant level was considered $P < 0.05$.

- *Psychometric Analysis*

(1) Floor and Ceiling effects:

Floor and Ceiling effects were calculated as the number and percentage of the total patients at the lowest and highest possible scores. The frequency of minimum and maximum scores should be less than 15% to ensure that Sf-36 is capturing the full range of possible responses in subjects with physical disabilities.

(2) Reliability:

- *Internal consistency*

Internal consistency was measured by Cronbach's α separately for each subscale. The cutoff 0.7 showed acceptable, 0.8 to 0.89 was good, and more than 0.9 represented excellent internal consistency[17].

- *Test-retest reliability*

Test-retest Reliability was measured by intra-class correlation (ICC) with 95% confidence intervals calculated using a two-way random effect model and "absolute agreement". In our study, ICC ≤ 0.40 is considered poor to fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 good agreement, and ≥ 0.80 excellent agreement[21]. A time interval of 1 week between test and re-test was considered long enough to prevent recall bias.

(3). Validity:

- *Convergent Validity*

Convergent validity is determined by scale-to-scale Pearson correlations between SF-36 and Moorong self-efficacy, with higher value (r ≥ 0.70) indicating the scales of different instruments measuring the same trait. A cut-off coefficient between relevant subscale was considered acceptable (r=0.3)[22].

- *Divergent Validity*

Correlation between subscales of HADS and subscales of SF-36 were conducted to explore divergent validity. We supposed that patients with a higher anxiety and depression scores had a worse health perception based on their SF-36 scores. We excepted a negative correlation between subscales and r=0.3 cut-off coefficient between subscales[22].

- *Concurrent Validity*

Evaluating concurrent validity, we investigated the Pearson correlation between mean score of MCS and PCS with their four theoretical subscales.

- *Discriminant validity*

Known-groups validity is demonstrated when a test or questionnaire can discriminate between two groups known to be different in terms of intended variable. Using the known groups method, a criterion is chosen that can identify the presence or absence of a particular characteristic. For this purpose, we randomly selected 120 normal cases among patients' families refereeing to the same clinics. We asked them to fill the SF-36 separately from their patients. We expected that the SF-36 could discriminate quality of life of individuals with physical disabilities from their normal family members. It was assumed that quality of life of normal family members was higher than the patients with physical disabilities.

- *Construct Validity*

To examine whether the 36 items of Sf-36 measured the same health related concepts in subjects with physical disabilities, we performed exploratory factor analyses using principal component analysis, with Varimax rotation. The Bartlett's test of sphericity was performed to evaluate the adequacy of sampling. Kaiser-Meyer-Olkin (KMO) test conducted to assess the adequacy of sample size. Factor loadings higher than 0.45 were considered to support factor construction and eigenvalue ≥ 1 was selected. We conducted Confirmatory Factor Analysis (CFA) among 300 other samples who were also physically disabled and they were recruited based on the convenient sampling method. The measurement model performed and seven theoretically predicted factors extracted.

We used the following indexes to adequate the fitness of the model: Chi-square (to evaluate the degree of difference between the empirical matrix and the theoretical model matrix, while

lower value of chi-square shows better adjustment), Normed fit index (NFI), Goodness-of-fit index (GFI), Comparative fit index (CFI), and Tucker-Lewis Index (TLI) (closer to 1.0 measure shows better model quality), Chi square/ degree of freedom (CMIN/DF.) is the minimum discrepancy divided by its degrees of freedom and the ratio should be close to 1 for better models. We calculated the root means square error of approximation (RMSEA) and values less than 0.08 were assumed acceptable[23].

Results

Sample characteristic and Statistical Analysis

A total number of 305 subjects with physical disabilities recruited, among which 191 (62.6%) were male and the mean age ± SD was 36.20 ±11.0. years (ranged from 18-86 yrs.). From all, 20% of participants had academic university certificate. Approximately more than 50% were married and 137 (50.0%) participants had severe disability. The mean duration of disability ±SD was 24.6 ± 14.5 years (ranged from 1-70 yrs.). Demographic characteristics of participants in sample 2 were similar to sample 1 as illustrated in Table 1.

Table 1. Demographic characteristics of participants of sample 1 for EFA and sample 2 for CFA

Characteristic		Sample number n (%) or MD±SD Sample 1 (*n*=305)	Sample number n (%) or MD±SD Sample 2 (*n*=300)
Age		36.20 ±11.00	37.10 ±12.00
sex	Male	191 (62.6)	196 (65.3)
	Female	114 (37.4)	104 (34.6)
Education	College study	63 (21)	57 (19)
	Illiterate	30 (9.8)	34 (11.3)
	Under diploma	212 (70)	209 (69.6)
Marital status	Single	136 (44.6)	124 (41.3)
	Married	156 (51.1)	162 (54.0)
	Separated	11 (3.6)	13 (4.3)
	Deceased spouse	2 (0.6)	1 (0.3)
Duration of disability (year)		24.62 ±14.48	22.52 ±13.19
Disability level	Low	90 (29.2)	92 (30.1)

	Moderate	182 (59.7)	180 (60.0)
	Severe	33(10.5)	28(9.3)
Physical activity	Low	2 (0.6)	7 (1.6)
	Moderate	163 (53.4)	143 (47.7)
	High	140 (45.9)	150 (50..0)
Current job	Unemployed	153 (51.0)	147(49.0)
	Worker	25 (8.0)	22 (7.3)
	Employee	18 (6.0)	15 (5.0)
	Retired	3 (1.00)	5 (1.6)
	Self-employed	74 (24.2)	79 (26.3)
	Housekeeper	32 (10.4)	32 (10.6)
Comorbidity	Yes	96 (32.0)	97 (32.3)
	No	209 (69.7)	203 (67.6)
Living place	City	294 (96.4)	288 (96.0)
	Village	11 (3.6)	12(4.0)

The most common physical disability among subjects were congenital abnormalities followed by injuries due to road accidents.

Psychometric analysis

(1) Floor and Ceiling effect

Role-physical item and Role emotional factors respectively had floor effect of 40.1% and 29.1%. Bodily pain, social functioning and Role-emotional items respectively had ceiling effect of 16.6%, 15.9% and 36.1%. All other items showed lower than 15% floor and ceiling effects. Floor and ceiling effects for subjects with physical disabilities are reported in Table 2 and items with >15% floor and ceiling effect is represented with bold font.

(2) Reliability

- *Internal consistency*

Internal consistency reliability is reported for subjects with physical disabilities in Table 2 and all eight theoretical subscales and both main domain (MCS and PCS) had greater than 0.7 Cronbach's α, with the 0.93 for the physical functioning.

- *Test-retest Reliability*

Test-retest reliability is reported via Intra-class correlation. As it is presented in Table 2 all the items showed excellent agreement (ICC ≥ 0.80) except vitality, Role-emotional and mental health items (ICC =0.79, 0.71, 0.72, respectively) that had good agreement.

Table 2. Internal consistency reliability, test-retest reliability, and floor/ceiling effect of the Sf-36 in sample 1(N =305).

	Internal consistency reliability	**Test-retest reliability**	**Floor and ceiling effects [n(%)]ǂ**	
	Cronbach's α	(Intra-class correlation, ICC)	Floor effect	Ceiling effect
Sf-36				
Physical functioning (PF)	0.93	0.83	21 (7)	10(3)
Role-physical (RP)	0.78	0.88	**121(40.1)**	18(6)
Bodily pain (BP)	0.84	0.81	8(2.6)	**50 (16.6)**
General health (GH)	0.70	0.82	2(0.7)	1(0.3)
Vitality (VT)	0.72	0.79	1(0.3)	5(1.7)
Social functioning (SF)	0.78	0.80	7(2.3)	**48 (15.9)**
Role-emotional (RE)	0.78	0.71	**88(29.1)**	**109(36.1)**
Mental health (MH)	0.82	0.72	1(0.3)	5(1.7)
Physical component summary (PCS)	0.79	0.85	0(0)	0(0)
Mental component summary (MCS)	0.78	0.83	0(0)	0(0)

ǂ Floor and ceiling effects [n (%)]: the numbers and percentages of the patients at the lowest and highest possible scores.

Floor effect or ceiling effect ≥ 15% was in bold.

(3) Validity

- *Convergent Validity:*

The correlation between both subscales of SF-36 and its two domains including PCS and MCS with Moorong self-efficacy total score as well as its two subscales including daily activity and social functioning were significant with correlation coefficients ranging from 0.22 to 0.56. The results are shown in Table 3. *(a)*

- *Divergent Validity:*

The level of adaptation between SF-36 and HADS was calculated. The strongest correlation was between emotional well-being subscale and HADS total score (Pearson r=-0.65) and vitality subscale with HADS total score (Pearson r= -0.65). The weakest correlation was between physical functioning subscale of SF-36 and Anxiety subscale of HADS with the correlation of -0.24.

The PCS domain of SF-36 had a weak correlation with HADS total score while the correlation between MCS domain of SF-36 with HADS total score was strong with the correlation of -0.61. Results are presented in Table 3. *(b)*

Table 3. The correlation between SF-36 and Moorong Self efficacy subscales *(a)* and level of correlation between SF-36 and HADS *(b)*

SF-36	Moorong self-efficacy subscale (*a*)		HADS total score and subscales*(b)*		
	Daily Activity	Social functioning	Total score	Depression subscale	Anxiety subscale
Physical functioning	0.46*	0.30*	-0.31*	-0.34*	-0.24*
Role physical	0.35*	0.25*	-0.41*	-0.36*	-0.38*
Role emotional	0.34*	0.28*	-0.46*	-0.35*	-0.47*
Vitality	0.53*	0.54*	-0.65*	-0.57*	-0.60*
Emotional well-being	0.42*	0.46*	-0.65*	-0.54*	-0.62*
Social functioning	0.48*	0.45*	-0.54*	-0.52*	-0.47*
Bodily pain	0.40*	0.39*	-0.52*	-0.45*	-0.48*
General health	0.49*	0.38*	-0.53*	-0.52*	-0.45*
PCS	0.43*	0.30*	-0.32*	-0.35*	-0.25*
MCS	0.37*	0.43*	-0.61*	-0.48*	-0.59*

PCS: physical component summary, MCS: mental component summary
* All *P* value were lower than 0.0001

- *Concurrent Validity*

Concurrent validity was evaluated between two main domains of SF-36 and their four related subscales. Table 4 shows moderate to high correlation coefficient from 0.49, for general health and PCS, to 0.88 for physical functioning and PCS.

Table 4. Concurrent Validity

SF-36 main domains	SF-36 physical and mental subscale			
PCS	Physical functioning	Role of physical	Bodily pain	General health
	0.88*	0.64*	0.64*	0.49*
MCS	Social functioning	Vitality	Mental health	Role emotional
	0.64*	0.78*	0.87*	0.73*
PCS: physical component summary, MCS: mental component summary * *P* value <0.0001				

- *Discriminant Validity*

We evaluated discriminant validity by comparing the SF-36 score in normal family members of physically disabled patients with disabled patients' SF-36 scores. In a further approach discriminant validity was calculated comparing score of both domains of SF-36 in different severity levels of physical disability. Results were presented in Table 5, showed that physical disabled patients were significantly different from their normal family members in terms of PCS scores. However, patients had a lower level of MCS, not significant, rather than their family members.

Table 5. Discriminant validity comparing physically disabled versus normal people in terms

SF36	***Normal family members*** (Mean±SD)	***physically disabled*** (Mean±SD)	***P value***	**Disability Level**		
				Moderate (Mean±SD)	**Severe** (Mean±SD)	*P* value
PCS	54.62 ± 8.22	36.06 ± 9.10	<0.0001*	38.18±9.21	33.19±8.15	<0.001*
MCS	53.42 ± 9.31	51.56±10.82	=0.20*	51.90±11.08	51.19±10.61	=0.70
PCS: physical component summary, MCS: mental component summary * *P* value <0.0001						

- *Construct Validity*

Explanatory factor analysis (EFA)

Factor analysis showed two principal components for Sf-36 that may represent physical and mental component summary. Results of Principal Component Analysis are represented in

table 2. Role Emotional and Mental Health factors are included in Mental component scale, while Physical Functioning, Role Physical, Bodily Pain and General Health are included in Physical component scale. Social Functioning belong to Physical component in a borderline condition and Vitality factor belongs to mental component.

The KMO measure of sampling adequacy showed a value of 0.91. The Bartlett's Test of Sphericity was 6153.24 with *p*-value <0.0001.

Principal Component Analysis also revealed seven factors. Considering Factor loadings larger than 0.45 and eigenvalues equal to 1, Items 29 and 31(from vitality) loaded on Factor 2(Mental Health), Items 26 and 30(Mental Health) loaded on Factor 3(Vitality), Items 13 and 14 (from Role Physical) loaded on Factor 4(Role Emotional), Items 20 and 32 (from Bodily Pain) and Items 1, 34 and 36 (from General Health) loaded on Factor 5(Social Functioning). In this way Bodily Pain factor merged with social functioning and changed number of factors from 8 to 7. Results of Principal Component Analysis for factors are represented on table 3. Factor analysis identified seven principal components that jointly accounted for 31.56% of the variance.

Table 6. Factor Analysis for component factors comparing to the original SF-36; each color represents eight original factors

SF-36 Subscales								
Original SF-36 Subscales		Factor 1 **Physical functioning**	Factor2 **Mental Health**	Factor 3 Vitality	Factor 4 **Role limitation due to emotional problems**	Factor 5 **Pain – Social functioning**	Factor 6 **General Health**	Factor 7 **Role limitation due to physical Health**
10	Walking **several blocks**	**0.87**						
9	Walking **more than a mile**	**0.86**						
11	Walking **one block**	**0.85**						
7	Climbing **one** flight of stairs	**0.84**						
6	Climbing **several** flights of stairs	**0.83**						
8	Bending, kneeling, or stooping	**0.81**						

4	**Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	0.72						
3	**Vigorous activities**, such as running, lifting heavy objects, participating in strenuous sports	0.66						
5	Lifting or carrying groceries	0.64						
12	Bathing or dressing yourself	0.43						
29	Did you feel worn out?		0.80					
31	Did you feel tired?		0.75					
25	Have you felt so down in the dumps that nothing could cheer you up?		0.69					
28	Have you felt downhearted and blue?		0.68					
24	Have you been a very nervous person?		0.68					
27	Did you have a lot of energy?			0.77				
30	Have you been a happy person?			0.76				
26	Have you felt calm and peaceful?			0.68				
23	Did you feel full of pep?			0.63				
17	Cut down the **amount of time** you spent on work or other activities as a result of any emotional problems				0.72			
18	**Accomplished less** than you would like as a result of any emotional problems				0.72			
19	Didn't do work or other activities as **carefully** as usual as a result of any emotional problems				0.67			
13	Cut down the **amount of time** you spent on work or other activities as a result of your physical health				0.57			
14	**Accomplished less** than you would like as a result of your physical health				0.54			
22	During the **past 4 weeks,** how much did **pain** interfere with your normal work (including both work outside the home and housework)?					0.70		
20	During the **past 4 weeks,** to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?					0.61		

36	My health is excellent					**0.59**		
21	How much **bodily** pain have you had during the **past 4 weeks**?					**0.59**		
34	I am as healthy as anybody I know					**0.56**		
32	During the **past 4 weeks**, how much of the time has **your physical health or emotional problems** interfered with your social activities (like visiting with friends, relatives, etc.)?					**0.51**		
1	In general, would you say your health is:					**0.45**		
35	I expect my health to get worse						**0.76**	
33	I seem to get sick a little easier than other people						**0.67**	
15	Were limited in the **kind** of work or other activities as a result of your physical health							**0.78**
16	Had **difficulty** performing the work or other activities (for example, it took extra effort) as a result of your physical health							**0.63**
% Variance		31.56	12.22	5.58	4.50	3.9	3.4	2.96
Egenvalue		11.05	4.28	1.95	1.58	1.35	1.90	1.04

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 9 iterations.

Confirmatory factor analysis (CFA)

The CFA conducted on the questionnaire to evaluate the fitness of the model obtained from the EFA. We also use the second-order factorial validity for the SF-36. While factor 1, factor 5, factor 5, and factor 7 loaded on Physical Factor (PCS), factor 2, factor 3, factor 4 were loaded on Mental Factor (MCS). Covariance matrixes were applied and fit indexes were calculated. Figure. 1 shows the best model fit. All fit indices were good, with chi-square ($\chi 2/DF$) was equal to 2.61 ($p < .001$). The RMSEA of the model was 0.07 (90% CI = 0.07–0.08), AGFI=0.75, GFI = 0.78, CFI= 0.85, and NFI = 0.78. As shown in Fig 1. Factor

loadings from 0.49 to 0.94 for PCS and ranged from 0.75 to 0.82 for MCS. All factor loadings were more than 0.60, with an exception of factor 4, which included a factor loading of 0.45 on item 13.

Discussion

The Persian version of SF-36 is a valid and reliable health survey for people who are physical disabled. We showed that the Persian SF-36 was appropriate for use among people with physical disability. We found that the Persian people with disability scored PCS of 36.10 and MCS of 51.6 out of 100. These scores are not surprising as the mean score of physical and mental dimensions in the Iranian general population ranged from 61.4 to 87.8 in the validation study of the Iranian version.

Internal consistency using Cronbach α was found to be good to excellent for all theoretical components and more than the cut off 0.70, which is consistent with what was reported for those in the general population[24]. All ICC coefficients for all components were greater than 0.80, with the exception of the Role emotional and Mental Health (ICC=0.71, 0.72, respectively), confirming acceptable internal consistency.

In this study, no ceiling and floor effects were found in both PCS and MCS. However, high floor effects were found in Role emotional and Role physical subscales, similarly as what has been observed in Ruta et al. (1998) study and Bunevicius et al. (2007), which were conducted among patients with coronary artery disease and Brian trauma[25,26]. The result may be due to the similarity in the physical and mental complication of the physical disability in coronary artery disease. While Bodily pain and Social functioning subscales were slightly higher than the accepted cut off 15%, we found a high ceiling effect for Role emotional. It might be related to patients' good therapeutic services leading to the high scoring they allocated to these subscales. As the original and Persian version of SF-36 haven't reported the floor and

ceiling effect values, future studies should provide modifications to the response items to eliminate the ceiling effect presented in the subscales.

In convergent validity the correlation between Moorong subscales and all SF-36 subscales and two domains were greater than the threshold of 0.3 with the exception of correlation between role emotional and social functioning (r=0.28) and role physical and social functioning (r=0.25). The greatest correlation was observed between vitality and Moorong two subscales (R=0.53 and 0.54 respectively). This is because physical health is represented directly in daily activities, which is also included in Moorong self-efficacy scale.

In divergent validity, the PCS domain of SF-36 had a weak correlation with HADS total score while there was a strong correlation between MCS domain of SF-36 and HADS total score. This strong correlation between these two components, MCS and HADS, may happen because of same nature of anxiety and mental component.

Concurrent validity was examined with level of association between SF-36 main domains and its respected subscales. We observed strong correlation between all domains with exception of general health and PCS (r=0.39). Concurrent validity shows that SF36 and its subscales are measuring same related concept and constructs.

Discriminant validity was assessed comparing two groups of subjects including people with disability and their family members. We also investigated discriminant validity between subjects with moderate and severe physical disabilities. We observed significant difference between PCS of subjects with physical disabilities and their family members. It is not surprising that no significant difference was observed between two groups. However, score of MCS was more than PCS. Family members achieved a greater but not significant MCS scores in comparison to subjects with physical disabilities. This shows that quality of life in subjects with physical disabilities is directly affected by their physical function, but mental health is not affected as much as Physical health. This pattern was also observed in the disability level.

Patients with severe disability had lower level of physical health. Studies has shown that subjects with physical disabilities grade lower on the physical functioning of the SF-36 than general populations [27,28]. Previous studies showed that SF-36 discriminate differences between groups with disabilities compared with the general population[29].

However, original SF-36 has eight factors. Results from explanatory factor analysis showed that in current study, all the items loaded on seven factors. First factor included 10 items, just like the original SF-36. Some of the factors belonging to second and third factors were exchanged between these two factors. This shows that mental health and vitality subscales points out to the same concepts in subjects with physical disabilities.

Mental health and vitality and vigorousness are a united concept for individuals with physical disabilities of this study. And this people feel more mentally healthy when they are vital and vigorous.

Items 13 and 14 that belong originally to the factor "Role limitation due to physical Health" are now added to "Role limitation due to emotional problems". Studies show that one's reduced physical capacity in interaction with the environment can impose on emotional and social role functioning (36) The merging of the two distinct components of limitation, physical and emotional, shows that due to lifelong disability, the physical disability has imposed heavy emotional loads on physical performance, in a way that now these limitations have a united concept for a person with physical disabilities .

Factor 5 and 6 of original SF-36 are the other two factors that are now merged into one factor including Bodily pain and social functioning. As the pain limits social functioning we can interfere that these two concepts are measuring one feature for subjects with physical disabilities.

In addition, studies show that Physical disability leads to physical and social consequence, such as inactive life style, losing of job, being slower performance, disruption in relationships among family members. These negative issues effect on both one's mental and physical

health[30]. So, in these population the meaning of pain, social functioning, and role limitation due to physical health were loaded on one factor, meaning that the having similar concept. Factor six of current study is exactly same as the factor 7 of original SF-36, which is evaluating General Health. The last factor, in the current study was Role limitation due to physical Health. Our EFA results showed that there were only seven factors underlying the SF-36 Persian version in subjects with physical disabilities. At the item level, they showed that half of the items of the VT scale (VT3 & VT4) merged into the MH factor, and the two items of the SF scale were not extracted as a distinct factor. One item (SF2) emerged as the MH factor; the other one (SF1) had no substantial loadings, but was regarded on the RE factor because its largest loading was on the RE factor. At the item level, they showed that half of the items of the VT scale (VT3 & VT4) merged into the MH factor, and the two items of the SF scale were not extracted as a distinct factor. One item (SF2) emerged as the MH factor; the other one (SF1) had no substantial loadings, but was regarded on the RE factor because its largest loading was on the RE factor. At the item level, they showed that half of the items of the VT scale (VT3 & VT4) merged into the MH factor, and the two items of the scale were not extracted as a distinct factor. One item (SF2) emerged as the MH factor; the other one (SF1) had no substantial loadings, but was regarded on the RE factor because its largest loading was on the RE factor. At the item level, they showed that half of the items of the VT scale (VT3 & VT4) merged into the MH factor, and the two items of the SF scale were not extracted as a distinct factor. One item (SF2) emerged as the MH factor; the other one (SF1) had no substantial loadings, but was regarded on the RE factor because its largest loading was on the RE factor. At the item level, they showed that half of the items of the VT scale (VT3 & VT4) merged into the MH factor, and the two items of the SF-36 scale were not extracted as a distinct factor. One item (SF2) emerged as the MH factor; the other one (SF1) had no substantial loadings, but was regarded on the RE factor because its largest loading was on the RE factor. The seven-factor structure was constant with the results in Wu et al.'s in

Taiwan and Thumboo et al.'s study in Singapore Thumboo et al.'s study in Singapore Thumboo et al's study in Singapore [31,32]. They result revealed that half of the items of the vitality subscales (29, 31) extracted into mental health subscales. Two items of mental health subscales (26, 30) loaded on vitality subscales. In two studies, the findings were the same as the EFA findings. This may suggest this seven-factor structure is valid across different countries and different samples. However, in contrast to the current study, Wu et al. concluded that EFA and CFA would not have the same result.

But both of the study conducted in East Asia and this may lead to make the same interpretation regarding to physical and mental health in SF-36, as we know health-related quality of life culturally constructed as people seek safety and security[33]. However, living with disability is a process of constant change and lead to change in physical, mental, and social health. These consequences may make a difference in the structure of the SF-36 subscales. Previous studies regarding validation of SF-36 in Asia shown that there is a mixture interpretation of the some mental and physical components[34]. Although Montazeri et al. (2005) validated the SF-36 with a first-order factorial analysis, we should make a caution using the SF-36 in persons with physical disability.

This study had some limitations. First, our study design was conducted in cross-sectional data, and further psychometric analyses such as responsiveness to clinical changes were not assessed. Second, we selected persons with physical disability from physiotherapy rehabilitation centers, and thus, the recruited patients may not be representative of typical Iranian population with all level of physical disability, as only one tenth of them were categorized as severe. We recommend that individuals with physical disabilities should be recruited from various communities beyond rehabilitation centers to include various subjects. More research is needed to evaluate if the current findings can be generalized to different population of patients with physical disabilities such as Parkinson, Multiple Sclerosis. Third, we didn't evaluate a physical activity questionnaire to understand the difference of QoL

scores across the level of physical activity. Because of using various questionnaires and Possibility of patient's fatigue, we didn't apply a specific quality of life tool. In that case, we could assess the correlation between SF-36 with a specific quality of life questionnaire like what was developed by Ebrahimzadeh et al. (2014)[35]. However, for people with a physical disability a questionnaire is needed that includes more population specific physical activity patterns. [36,37]

In conclusion, the SF-36 is a reliable and valid tool in physical disables. However, SF-36 show insufficient eight-factorial validity. The confirmatory factor analysis confirmed the result of seven-factor extracted from explanatory factor analysis. The CFA was the same as the EFA results and may suggest that this seven-factor structure should be further assessed by other individuals with physical disabilities for cross-validation. The differences in the SF-36 scores, between subjects with physical disabilities and general population as well as among the level of disability confirm that SF-36 can be used in theses group, and then measuring outcome of rehabilitation services and managing this population.

Acknowledgments

We greatly thank all persons who participate in this study. We also thank for Student Research Committee, Kermanshah University of Medical Sciences, Kermanshah, Iran. This study was supported and approved by ethic committee of Vice-Chancellery for Research and Technology (Code: IR.KUMS.REC.1398.899).

Authors' contributions

All authors contributed to the study conception and design. All authors read and approved the final manuscript. Detailed contribution of authors follows. Nassim Ahmadi contributed proposal writing, statistical analysis, interpretation of data sand article writing. Mojgan Rajati contributed in interpretation of data and article writing. Maryam Khazaei Pool contributed in

sampling, and article writing. Leili Salehi contributed in sampling, and article writing. Mohammad Ebrahimi participated in sampling and entering questionnaire data into SPSS software. Fatemeh Rajati was the Supervision of research, and she managed and participated in whole process of proposal writing, sampling, statistical analysis, interpretation of data, original draft preparation, and article writing.

Conflicts of interest

The authors declare no conflicts of interest in relation to this article.

Reference

1. Spencer L James, D. A., Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi et al. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet, 392*(10159), 1789-1858, doi:https://doi.org/10.1016/s0140-6736(18)32279-7.
2. Kuvalekar, K., Kamath, R., Ashok, L., Shetty, B., Mayya, S., & Chandrasekaran, V. (2015). Quality of Life among Persons with Physical Disability in Udupi Taluk: A Cross Sectional Study. *Journal of Family Medicine and Primary Care, 4*(1), 69-73, doi:https://doi.org/10.4103/2249-4863.152258.
3. Walters, S. J., Munro, J. F., & Brazier, J. E. (2001). Using the SF-36 with older adults: a cross-sectional community-based survey. *Age and Ageing, 30*(4), 337-343, doi:https://doi.org/10.1093/ageing/30.4.337.
4. Patrick, D. L., & Deyo, R. A. (1989). Generic and disease-specific measures in assessing health status and quality of life. *Med Care, 27*(3 Suppl), S217-232, doi:https://doi.org/10.1097/00005650-198903001-00018.
5. Moen, V. P., Eide, G. E., Drageset, J., & Gjesdal, S. (2019). Sense of Coherence, Disability, and Health-Related Quality of Life: A Cross-Sectional Study of Rehabilitation Patients in Norway. *Archives of Physical Medicine and Rehabilitation, 100*(3), 448-457, doi:https://doi.org/10.1016/j.apmr.2018.06.009.
6. Alkan, B. M., Fidan, F., Tosun, A., & Ardıçoğlu, Ö. (2014). Quality of life and self-reported disability in patients with knee osteoarthritis. *Modern Rheumatology, 24*(1), 166-171, doi:https://doi.org/10.3109/14397595.2013.854046.
7. Feroz, A. H. M., Islam, M. N., ten Klooster, P. M., Hasan, M., Rasker, J. J., & Haq, S. A. (2012). The Bengali Short Form-36 was acceptable, reliable, and valid in patients with rheumatoid arthritis. *Journal of Clinical Epidemiology, 65*(11), 1227-1235, doi:https://doi.org/10.1016/j.jclinepi.2012.05.004.
8. Ko, S., Choi, W., Chae, S., Kwon, J., & Lee, Y. (2018). Correlation between Short-Form 36 Scores and Neck Disability Index in Patients Undergoing Anterior Cervical Discectomy and Fusion. *Asian spine journal, 12*(4), 691-696, doi:https://doi.org/10.31616/asj.2018.12.4.691.
9. Montazeri, A., Goshtasebi, A., Vahdaninia, M., & Gandek, B. (2005). The Short Form Health Survey (SF-36): Translation and validation study of the Iranian version.

Quality of Life Research, 14(3), 875-882, doi:https://doi.org/10.1007/s11136-004-1014-5.

10. Riazi, A., Hobart, J. C., Lamping, D. L., Fitzpatrick, R., Freeman, J. A., Jenkinson, C., et al. (2003). Using the SF-36 measure to compare the health impact of multiple sclerosis and Parkinson's disease with normal population health profiles. *J Neurol Neurosurg Psychiatry, 74*(6), 710-714, doi:https://doi.org/10.1136/jnnp.74.6.710.
11. Bentsen, S. B., Rokne, B., & Wahl, A. K. (2013). Comparison of health-related quality of life between patients with chronic obstructive pulmonary disease and the general population. *Scand J Caring Sci, 27*(4), 905-912, doi:https://doi.org/10.1111/scs.12002.
12. Horowitz, A., Reinhardt, J. P., Boerner, K., & Travis, L. A. (2003). The influence of health, social support quality and rehabilitation on depression among disabled elders. *Aging Ment Health, 7*(5), 342-350, doi:https://doi.org/10.1080/1360786031000150739.
13. Chang, K.-H., Liao, H.-F., Yen, C.-F., Hwang, A.-W., Chi, W.-C., Escorpizo, R., et al. (2015). Association between muscle power impairment and WHODAS 2.0 in older adults with physical disability in Taiwan. *Disability and Rehabilitation, 37*(8), 712-720, doi:https://doi.org/10.3109/09638288.2014.940428.
14. Johansen, H., Bathen, T., Andersen, L. Ø., Rand-Hendriksen, S., & Østlie, K. (2018). Chronic pain and fatigue in adults with congenital unilateral upper limb deficiency in Norway. A cross-sectional study. *PLOS ONE, 13*(1), e0190567, doi:https://doi.org/10.1371/journal.pone.0190567.
15. Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care, 30*(6), 473-483, doi:https://doi.org/10.1097/00005650-199206000-00002.
16. Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatr Scand, 67*(6), 361-370, doi:https://doi.org/10.1111/j.1600-0447.1983.tb09716.x.
17. Nunnally, J., Jum, N., Bernstein, I. H., & Bernstein, I. (1994). *Psychometric Theory* (Vol. no. 972): McGraw-Hill Companies,Incorporated.
18. Montazeri, A., Vahdaninia, M., Ebrahimi, M., & Jarvandi, S. (2003). The Hospital Anxiety and Depression Scale (HADS): translation and validation study of the Iranian version. *Health and Quality of Life Outcomes, 1*, 14-14, doi:https://doi.org/10.1186/1477-7525-1-14.
19. Middleton, J. W., Tate, R. L., & Geraghty, T. J. (2003). Self-Efficacy and Spinal Cord Injury: Psychometric Properties of a New Scale. *Rehabilitation Psychology, 48*(4), 281-288, doi: https://doi.org/10.1037/0090-5550.48.4.281.
20. Rajati, F., Ghanbari, M., Hasandokht, T., Hosseini, S. Y., Akbarzadeh, R., & Ashtarian, H. (2017). Persian version of the Moorong Self-Efficacy Scale: psychometric study among subjects with physical disability. *Disabil Rehabil, 39*(23), 2436-2445, doi:https://doi.org/10.1080/09638288.2016.1226404.
21. Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological reports, 19*(1), 3-11, doi: https://doi.org/10.2466/pr0.1966.19.1.3.
22. Portney, L. G., & Watkins, M. P. (2015). *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River: Prentice Hall Health.
23. Furr, R. M. (2011). Scale Construction and Psychometrics for Social and Personality Psychology. doi: https://dx.doi.org/10.4135/9781446287866.
24. Garratt, A. M., & Stavem, K. (2017). Measurement properties and normative data for the Norwegian SF-36: results from a general population survey. *Health and Quality of Life Outcomes, 15*(1), 51, doi:https://doi.org/10.1186/s12955-017-0625-9.
25. Ruta, D. A., Hurst, N. P., Kind, P., Hunter, M., & Stubbings, A. (1998). Measuring health status in British patients with rheumatoid arthritis: reliability, validity and

responsiveness of the short form 36-item health survey (SF-36). *Br J Rheumatol, 37*(4), 425-436, doi:https://doi.org/10.1093/rheumatology/37.4.425.
26. Bunevicius, A. (2017). Reliability and validity of the SF-36 Health Survey Questionnaire in patients with brain tumors: a cross-sectional study. *Health and Quality of Life Outcomes, 15*(1), 92, doi:https://doi.org/10.1186/s12955-017-0665-1.
27. Andresen, E. M., Fouts, B. S., Romeis, J. C., & Brownson, C. A. (1999). Performance of health-related quality-of-life instruments in a spinal cord injured population. *Arch Phys Med Rehabil, 80*(8), 877-884, doi:https://doi.org/10.1016/s0003-9993(99)90077-1.
28. Vickrey, B. G., Hays, R. D., Harooni, R., Myers, L. W., & Ellison, G. W. (1995). A health-related quality of life measure for multiple sclerosis. *Quality of Life Research, 4*(3), 187-206, doi:https://doi.org/10.1007/BF02260859.
29. Ware, J. E., Kosinski, M., & Keller, S. (1994). *SF-36 physical and mental health summary scales: a user's manual*. Boston Health Assessment Lab.
30. Andersson, A.-L., Dahlbäck, L.-O., & Bunketorp, O. (2005). Psychosocial aspects of road traffic trauma–benefits of an early intervention? *Injury, 36*(8), 917-926, doi:https://doi.org/10.1016/j.injury.2004.09.019.
31. Thumboo, J., Wu, Y., Tai, E. S., Gandek, B., Lee, J., Ma, S., et al. (2013). Reliability and validity of the English (Singapore) and Chinese (Singapore) versions of the Short-Form 36 version 2 in a multi-ethnic urban Asian population in Singapore. *Quality of Life Research, 22*(9), 2501-2508, doi:https://doi.org/10.1007/s11136-013-0381-1.
32. Wu, C.-h., Lee, K.-l., & Yao, G. (2007). Examining the hierarchical factor structure of the SF-36 Taiwan version by exploratory and confirmatory factor analysis. *Journal of Evaluation in Clinical Practice, 13*(6), 889-900, doi:https://doi.org/10.1111/j.1365-2753.2006.00767.x.
33. Kagawa-Singer, M., Padilla, G. V., & Ashing-Giwa, K. (2010). Health-related quality of life and culture. *Semin Oncol Nurs, 26*(1), 59-67, doi:https://doi.org/10.1016/j.soncn.2009.11.008.
34. Tseng, H.-M., Lu, J.-f. R., & Gandek, B. (2003). Cultural Issues in Using the SF-36 Health Survey in Asia: Results from Taiwan. *Health and Quality of Life Outcomes, 1*(1), 72, doi:https://doi.org/10.1186/1477-7525-1-72.
35. Ebrahimzadeh, M. H., Soltani-Moghaddas, S. H., Birjandinejad, A., Omidi-Kashani, F., & Bozorgnia, S. (2014). Quality of life among veterans with chronic spinal cord injury and related variables. *Archives of trauma research, 3*(2). doi:https://doi.org/10.5812/atr.17917
36. van der Ploeg, H. P., Streppel, K. R., van der Beek, A. J., van der Woude, L. H., Vollenbroek-Hutten, M., & van Mechelen, W. (2007). The Physical Activity Scale for Individuals with Physical Disabilities: test-retest reliability and comparison with an accelerometer. *J Phys Act Health, 4*(1), 96-100, doi:https://doi.org/10.1123/jpah.4.1.96.
37. Washburn, R. A., Zhu, W., McAuley, E., Frogley, M., & Figoni, S. F. (2002). The physical activity scale for individuals with physical disabilities: development and evaluation. *Arch Phys Med Rehabil, 83*(2), 193-200, doi:https://doi.org/10.1053/apmr.2002.27467.