Using Confirmatory Factor Analysis (CFA) to Assess Structure of SF-36 (Persian Version) questions for normal people and patients with chronic diseases

Shideh Rafaati [1]   Zahra Bagheri [2]   Peyman Jafari [3]   Nahid Soltani [4]

PhD Student in Biostatistical Social Departerment in Health Promotion Research Center [1] , MSc of Educational Management [4] , Hormozgan University of Medical Sciences, Bandar Abbas, Iran. Assistant Professor Department of Biostatistics [2] , Associate Professor Department of Biostatistics [3] , Shiraz Universiuty of Medical Sciences, Shiraz, Iran.

Original Article

Abstract

Correspondence:
Peyman Jafari, PhD.
Department of Biostatistics,
Shiraz University of Medical
Sciences.
Shiraz, Iran
Tel:+98 9123172566
Email:
pjbiostat@gmail.com

Introduction: This study aims at analyzing factor structure of Short Form Healthy Survey (SF-36), Persian Version, for both normal people and patients with chronic diseases.

Methods: In this cross-sectional study, a total of 585 patients with chronic diseases and 642 over 18-years old normal people were selected using simple casual sampling method. When the survey was completed by the participants, the Cronbach's alpha coefficient and confirmatory factor analysis (CFA) were used to determine the survey's validity and reliability. T-test also was used to compare the average score of the survey's domains in both groups.

Results: For the group of patients with the chronic diseases, the highest mean belonged to the social functioning domain's score (43.11 ± 18.48). In contrast, for the normal people's group, vitality found the highest mean score (50.83 ± 9.95). T-test results showed that the mean difference in all domains except the role limitations because of social impediments for both normal people and patients was significant. Cronbach's alpha coefficient for all eight domains of the survey for the normal people varied from 0.68 to 0.90, which it indicated an acceptable reliability for the survey; whereas the coefficient for the eight domains of the survey for patients varied from 0.66 to 0.91. CFA results using goodness-of-fit indicators showed that the 8-factor structure, suggested by the developer, is not verified for the patients group and the CFA results in the normal people's group showed that the 8-factor model has an acceptable fitness for the data.

Conclusion: CFA is an effective method to analyze structure of different domains of SF-36 in various groups. For this reason, researchers need to be wary when they use the survey and also they need to check initially its reliability and validity in different sub-groups.

Key words: Factor Analysis - Quality of Life - Iran

Introduction:

When a measuring tool is translated from a language to another, its properties and measuring quality should be assessed in terms of validity and reliability. SF-6 questionnaire is one of the most popular questionnaires used in many countries including Iran to assess quality of life in normal people and patients with chronic diseases (1). What is very important in all surveys designed to appraise quality of life is that the questions which have been considered to assess only a single domain should just assess that domain; it means that the one-dimensionality condition must be met here. Likewise, the participant's answer to a certain question on a domain should not be influenced by his/her answers to other questions on the same domain (2). The studies on this area have suggested that any deviation from the mentioned rules can affect considerably the results of both reliability and validity tests. Accordingly, marginalizing the one-dimensionality condition upon designing a questionnaire may result in misguiding results according to which a proper question is included in the questionnaire (3,4). It is necessary to note that an incorrect assessment of quality of life may result in wrong estimations and decisions on offering the most proper medical methods and improving the health level. Factor analysis is a popular technique to determine validity and to analyze the factor structure of questionnaires. Generally speaking factor analysis is used in two cases. Exploratory factor analysis (EFA) is used when the researcher has no idea about the structure of relationships among the questions; but if the questions of each domain have been identified already, you may need to use confirmatory factor analysis (CFA) (5). Pakpour showed that CFA is able to differentiate patients in terms of their clinical and demographic variables (8). Keller et al. used CFA to assess reliability of SF-36 survey in ten countries including Denmark, Germany, Italy, Netherlands, Norway, Sweden, UK and the United States. In fact, this study aimed at analyzing whether the survey is sufficiently valid even after being translated into different languages? The results showed that only mental health and physical functioning domains of the survey have been interpreted in a similar way in these countries. In other domains, because of translation [limitations] and perhaps wrong interpretations, people from different nationalities had reached different perceptions about questions, which it in turn resulted in different findings (7). Regarding the importance of using a valid and reliable measuring tool, our purpose in this study is analyzing reliability and factor structure of SF-36, Persian Version, in both normal people and patients with chronic diseases.

Methods:

It is a cross-sectional and applied study. The factor structure of SF-36 was studied among over 18-years old people living in Shiraz City through two groups, normal people and patients with chronic diseases. A total of 585 patients with chronic diseases, including patients with diabetes (180), patients with AIDS (100), patients who need dialysis (150) and patients with liver transplant (155), who had referred to Shiraz City hospitals for treatment were selected as the group of patients with chronic diseases. A total of 642 over 18-years old normal people were selected as the group of normal people through simple casual sampling method. The normal people group members were selected as follows: After taking necessary permissions from Education Organization of Shiraz City and receiving names and addresses of public male and female schools (i.e. primary, guidance and high schools), Shiraz City was divided into four geographical zones in order to leverage cultural, economic and social differences. After justifying students, researchers asked them to fill SF-36 survey by one of their parents provided that the parent who is asked to fill the survey should be free from chronic disease of AIDS, diabetes, dialysis and liver transplant surgery experience. Most parents who have student children are middle-aged; hence, the author asked a number of retirees of the Retirement Association and university students to fill the survey in order to find normal members from all age groups of over 18-years old.

SF-36 survey is one of the most popular surveys used to assess the quality of life in normal people and patients with chronic diseases in many countries across the world including Iran. Both validity and reliability of the survey have been examined through several studies. The important variables in this study are all questions and domains. The

survey has been designed with 36 questions in eight domains:

Physical Functioning (PF) with 10 questions; role limitations because of physical impediments (RP) with 4 questions; body pain (BP) with 2 questions; general health perceptions (GH) with 5 questions; vitality (VT) with 4 questions; social functioning (SF) with 2 questions; role limitations because of emotional impediments (RE) with 3 questions; and mental health (MH) with 5 questions.

Likert Scale method was considered to answer the questions. All 10 questions that constitute the physical functioning domain have three-point answers (yes, I have many impediments, yes, I have some impediments, No, I have not any impediment). Answers of other questions, except question #7, which was about the body pain domain, are in the form of 5 point. Options of question #7 have been designed in six points (none, very weak, mild, average, strong and very strong. It is necessary to note that some questions are reversed, but finally all questions are converted to the positive mode according to the instruction included in SF-36 and scores, in each case of 8 domains, was calculated from 0 to 100 through adding the interviewee's selected answers for each question to http://www.sf36.com/demos/sf-36v2. Higher scores represent high quality of life.

Data were described after handling with SPSS. Cronbach's alpha was used to analyze the survey's reliability and CFA was used to study the factor structure. LISREL software (ver. 8.54) was employed to carry our confirmatory factor analysis procedure.

The one-dimensionality condition indeed means that all questions in a domain just assess that domain. The condition is checked with CFA. The goodness of fit test is used to analyze the one-dimensionality condition. If the model is completely fitted with the data, then the condition was met; it means that all question of a domain only assess that dimension. The fitness of the fitted model is analyzed using chi-square, RMSEA, NNFI, CFI, PMR. Values smaller than chi-square value represent model's fitness. According to some studies, it has been suggested that for accepting the model based on the chi-square, degree of freedom needs to be less than 3. $NNFI > 0.92$, $CFI > 0.9$, and $PMR = 0$ represent the acceptable fitness of the model (7, 8, and 9).

Results:

Cronbach's alpha coefficient was used to analyze reliability of SF-36. It was 0.93 for normal people and varied from 0.68 to 0.90 for micro-scales. Cronbach's alpha coefficient of the questionnaire for the patients group measured 0.92 which demonstrated the relatively high reliability of SF-36 in both groups.

Table 1. Frequency distribution by gender in each group

	Women		Men		Total	
	Number	Percent	Number	Percent	Number	Percent
Normal group	394	61.4	248	38.6	642	52.3
Patient group	249	42.6	336	57.4	585	47.7

Table 2. Mean value and SD of people's age by gender in each group

	Women	Men
	SD ± average age	SD ± average age
Normal group	41.96±1.07	40.83±1.11
Patient group	49.18±1.12	49.38±1.22

Regarding p-value, it is seen that the mean difference in all domains, except, role limitation because of emotional impediments, is significant.

Goodness-of-fit indicators values for the CFA (table 5) indicate that the one-dimensionality condition has been met for general health and limitation because of emotional impediments in both normal and patient groups. On other words, the predefined structure for each domain was verified. Likewise, limitation because of physical impediments in normal group was found a one-dimensional domain; it means that, as it had been guessed, considering 4 questions to measure this domain was confirmed.

Table 3. Mean value and SD of scores of questionnaire's domains and Cronbach's alpha coefficient for both groups

Domains	No. of questions	SD ± Mean	SD Error	Cronbach's alpha coefficient
Pationts Group				
Physical functioning	10	41.39 ± 16.40	1.36	0.91
Role limitations because of physical impediments	4	38.71 ±16.25	1.35	0.83
Vitality	4	41.71 ± 16.08	1.34	0.71
Mental health	5	38.82 ± 18.62	1.55	0.81
Social functioning	2	43.11 ± 18.48	1.54	0.74
Role limitation because of emotional impediments	3	38.62 ± 17.75	1.47	0.84
Body pain	2	42.17 ± 17.24	1.43	0.73
Normal Group				
Physical functioning	10	47.07 ± 10.19	0.84	0.90
Role limitations because of physical impediments	4	42.63 ± 9.11	0.76	0.84
Vitality	4	50.83 ± 9.95	0.83	0.73
Mental health	5	41.23 ± 11.11	0.92	0.75
Social functioning	2	45.05 ± 10.16	0.84	0.68
Role limitation because of emotional impediments	3	39. 60 ± 11.33	0.94	0.82
Body pain	2	48.39 ± 10.11	0.84	0.74
General health perceptions	5	45.92 ± 10.65	0.88	0.73
Physical functioning	10	47.07 ± 10.19	0.84	0.90

Table 4. **Comparing the average score of domains in normal and patient group using T-test for both groups**

Normal/patient group	Mean difference	t-value	Degree of freedom	P-value
Physical functioning	5.68	7.35	1225	0
Role limitations because of physical impediments	3.93	5.28	1225	0
Vitality	9.12	12.05	1225	0
Mental health	2.4	2.76	1224	0.006
Social functioning	1.93	2.3	1225	0.022
Role limitation because of emotional impediments	0.98	1.16	1225	0.243
Body pain	6.22	7.79	1225	0
General health perceptions	8.66	12.32	1222	0

Table 5. Goodness-of-fit indicators for CCFA model in eight domains of SF-36

	Chi-square	d.f.	RMSEA	NNFI	RMR	CFI
Physical functioning	436.93	35	0.14	0.97	0.086	0.98
limitation because of physical impediments	13.79	2	0.1	0.99	0.025	1
General health perception	24.07	5	0.081	0.98	0.039	0.99
Vitality	31.24	2	0.16	0.91	0.065	0.97
Limitation because of emotional impediments	0	0	0	The model is saturated, the fit is perfect		
Mental health	31.52	5	0.096	0.99	0.041	0.99
Physical functioning	542.53	35	0.15	0.98	0.1	0.98
limitation because of physical impediments	3.45	2	0.034	1	0.012	1
General health perception	26.44	5	0.082	0.98	0.041	0.99
Vitality	76.94	2	0.24	0.79	0.095	0.93
Limitation because of emotional impediments	0	0	0	The model is saturated, the fit is perfect		
Mental health	118.10	5	0.19	0.89	0.093	0.94

Table 6. Goodness-of-fit indicators after improving each domain

	Chi-square	d.f.	RMSEA	NNFI	RMR	CFI
	Patients Group					
Physical functioning	150.67	30	0.084	0.98	0.073	0.99
limitation because of physical impediments	0.28	1	0	1	0.0032	1
Vitality	0.0078	1	0	1.01	0.0007	1
Mental health	5.92	4	0.029	1	0.018	1
	Normal Group					
Physical functioning	85.66	30	0.054	0.98	0.082	0.99
Vitality	0	0	0	0.96	0.075	0.96
Mental health	10.83	2	0.082	0.99	0.024	1

RMSEA > 0.1 values suggest that it would be impossible to speak about meeting one-dimensionality condition of physical functioning, vitality and mental functionality in both groups and limitation because of physical impediment in patients group. Considering the correlation between pair questions which apply the most adjustment in chi-square value in each domain, the following table can be offered in order to improve the model's fit indicators.

Note: RMSEA: root mean square error of approximation, NNFI: non – normed fit index

CFI: comparative fit index, RMR: root mean square residual.

Conclusion:

In this study, confirmatory factor analysis made it clear that the 8-factor structure suggested by the developed is confirmed for the normal people group. In contrast, five questions on the mental health was not confirmed for the patients group; because the correlation between two pairs of the questions was so strong that the mental health was divided into two domains and it will be problematic, because in this study, the questionnaire will measure 9 domains, rather 8 domains, for the patients. In other words, the CFA results in patients group represent that the 5-factor model has not an acceptable fitness on data. In physical functioning, for patients with chronic diseases, correlation was found between questions 4 and 5, 7 and 8, 1 and 7. In this study, correlation between pair questions of all domains, except mental health for patients group, was not very strong that can question the one-dimensionality condition. However, a less strong correlation was evident among questions in some domains of the questionnaire for both normal

people and patients. For physical functioning domain of normal people, the strongest correlation was found between question # 1) Does your health condition limit you in doing heavy exercises, such as running? And question # 4) Does your health limit you in climbing stairs? But this correlation does not question the one-dimensionality condition. The CFA results also showed that in mental health domain of patients there is correlation between two questions, have you felt convenience and satisfaction within last month? and have you been happy within last month? SF-36 was used in a study (1993) on quality of life of four groups of patients in comparison to normal people and the results showed that 8 sub-scales in patients group is not confirmed (8) and our results are consistent with these results. CFA was introduced as a good method to asses SF-36 survey in a study (2013) conducted in Australia over 3014 people. Its results showed that there was a correlation between mental health and role limitation because of physical impediments (9). However, the results of a study (2012) conducted in Canada in various groups, in terms of gender and race, with the aim of assessing SF-36 survey, the CFA results showed that gender and race have no impact on people's general understanding of questions; it means that the questionnaire's reliability was similar in various groups (1). Pakpour et al. assessed the quality life of 144 hemodialysis patients in Iran using SF-36. The CFA results indicated that the survey can differentiate patients in terms of clinical and demographic variables (10). Keller et al. (1998) used CFA to assess reliability of SF-36 in ten countries including Denmark, Germany, France, Italy, Netherlands, Norway, Spain, Sweden, UK and the United States. In fact, this study aimed at analyzing whether the survey is valid sufficiently even after being translated into various languages? The results showed that only mental health and physical functioning domains of the survey have been interpreted equally in these countries. In other domains, because of translation and perhaps wrong interpretations, people from different nationalities had reached different perception of questions which in turn it resulted in different findings (11). Field et al. (2000) analyzed validity and reliability of SF-36 among 185 patients with cardiac disease and normal people using CFA and Cronbach's alpha. Their

study showed that SF-36 is a useful scale to assess quality of life in patients with cardiac diseases (12). A study (2013) in Singapore was conducted to assess English and Chinese versions of SF-36; the results suggested that the survey's validity was not the same in both groups (13). A study (2005) was conducted on people with stroke and control group. The CFA results showed that SF-36 is able to differentiate the two groups (14). A study (2007) in the United States analyzed reliability and factor structure of SF-36 among old people with the traumatic brain injury. In this study, CFA was used in order to examine the impact of surgery intensity and racial groups on how this survey works. The results, 8-domain structure of the survey were confirmed. Although the surgery intense did not affect the validity of survey, the racial groups were effective. After controlling the effect of racial diversity, the survey's validity for measuring the score of quality of life in old people with the traumatic brain injury was reported acceptable (2). Comparing results of previous studies and this study implies that certain factors such as the sample, geographical zone, gender, race, language, culture, etc. can affect the difference between results I various studies.

Confirmatory factor analysis is an effective method to analyze structure of domains of SF-36 across different groups. Our results indicated that considering five questions in the mental health domain of patients is denied, because the correlation between questions 1 and 3 was so strong that these they measured a separate domain independently. For this reason, researchers need to be wary when they use the SF-36 and also they need to check initially its reliability and validity in different sub-groups.

Limitations of the study:
The patients group only included patients with diabetes, AIDS, dialysis patients, and patients who had liver transplants. Since the group is just composed of patients with the mentioned diseases, then the results cannot be generalized to all patients with the chronic diseases. On the other hand, since the study has been conducted in Shiraz City, then its results cannot be generalized to whole Iran.

Recommendation:

Thanks to its broad application, it is recommended that its properties and factor measuring quality is studied in other cities and ethnic, lingual, cultural and groups in order to achieve a standard, valid and reliable tool on quality of life based on the results of such studies.

Acknowledgement:

References:

1. Lix LM, Osman BA, Adachi JD, Towheed T, Hopman W, Davison KS, et al. Measurement equivalence of the SF-36 in the canadian multicentre osteoporosis study. *Health Qual Life Outcomes*. 2012:10-29.

2. Yang C, Selassie AW, Carter RE, Tilley BC. Measuring QoL with SF-36 in Older Americans with TBI. *Applied Research in Quality of Life*; 2012;7:63-81.

3. Wolinsky FD, Miller DK, Andresen EM, Malmstrom TK, Miller JP. Health-related quality of life in middle-aged African Americans. *The Journals of Gerontology*. 2004;59:118-123.

4. Ali-Mohammadpour R, Yousefi Z. Factor analysis of SF-36 Persian version health-related quality of life questionnaire in Iran. *World Applied Sciences Journal*. 2008;3:548-554.

5. Montazeri A, Vahdaninia M, Mousavi SJ, Omidvari S. The Iranian version of 12-item Short Form Health Survey (SF-12): factor structure, internal consistency and construct validity. *BMC Public Health*. 2009;9:341-351.

6. Pakour AH, Nourozi S, Molsted S ,Harrison AP, Nourozi K, Fridlund B. Validity and reliability of short-form: 12 questionnaire in Iranian hemodialysis patients. *Iran J Kidney Dis*. 2011;5:175-781.

7. Keller SD, Ware JE Jr, Bentler PM, Aaronson NK, Alonso J, Apolone G, et al. Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*. 1998;51:1179-1188.

8. Meng H, King-Kallimanis BL, Gum A, Wamsley B. Measurement bias of the SF-36 Health Survey in older adults with chronic conditions. *Qual Life Res*. 2013:22:2359-2369.

9. Jafari P, Bagheri Z, Ayatollahi S, Soltani Z. Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQLTM 4.0 Generic Core Scales in school children. *Health Qual Life Outcomes*. 2012;10:27-38.

10. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociological Methods & Research*. 1992;21:230-358.

11. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107:238-246.

12. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ*. 1993;306:1440-1444.

13. Thumboo J, Wu Y, Tai E-S, Gandek B, Lee J, Ma S, et al. 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. *Qual Life Res*. 2013:22:2501-2508.

14. Ahmed S, Mayo NE, Corbiere M, Wood-Dauphinee S, Hanley J, Cohen R. Change in quality of life of people with stroke over time: True change or response shift? *Qual Life Res*. 2005;14:611-627.