

Research Paper

International Journal of Medical Sciences

2012; 9(7):521-526. doi: 10.7150/ijms.4503

The 36-Item Short Form Health Survey: Reliability and Validity in Chinese Medical Students

Yang Zhang¹, Bo Qu^{1⊠}, Shi-si Lun¹, Ying Guo² and Jie Liu³

1. The Research Centre for Medical Education, China Medical University, 92 North Second Road, Heping District, Shenyang 110001, China;

2. Faculty of Fourth Affiliated Hospital, China Medical University, 4 Congshan east Road, Huanggu District, Shenyang 110032, China;

3. Faculty of Health Statistics, School of Public Health, China Medical University, 92 North Second Road, Heping District, Shenyang 110001, China.

🖂 Corresponding author: E-mail: qubo6666@163.com; Tel.: +86-24-23256666-6049; Fax: +86-24-23261090.

© Ivyspring International Publisher. This is an open-access article distributed under the terms of the Creative Commons License (http://creativecommons.org/ licenses/by-nc-nd/3.0/). Reproduction is permitted for personal, noncommercial use, provided that the article is in whole, unmodified, and properly cited.

Received: 2012.04.21; Accepted: 2012.08.19; Published: 2012.08.27

Abstract

Objective: The 36-Item Short Form Health Survey (SF-36) is widely validated and popularly used in assessing the subjective quality of life (QOL) of patients and the general public. The aim of the study is to assess the psychometric properties of the 36-Item Short Form Health Survey (SF-36) in medical students in mainland of China.

Methods: The reliability and validity of the 36-Item Short Form Health Survey (SF-36) questionnaire were assessed by conducting a cross-sectional study of Chinese medical students in December 2011. All 1358 3rd year and 4th year medical students from 46 classes at China Medical University were investigated.

Results: The overall Cronbach's α coefficient of the SF-36 questionnaire was 0.791, while the respective Cronbach's α coefficients for each of the seven dimensions were > 0.70, except where the social function dimension was 0.631. Results showed that the SF-36 questionnaire was reliable and valid.

Conclusion: In general, this study provides evidence that the SF-36 questionnaire is suitable measures for assess the QOL of medical students in China.

Key words: Medical student; Quality of life; The 36-item short form health survey (SF-36); Reliability; Validity.

Introduction

The 36-Item Short Form Health Survey (SF-36) questionnaire was developed by the Boston Health Research Institute in the United States. The SF-36 questionnaire provides a concise method that is mainly used to check the health status of members of the general population aged 14 years or over¹. The SF-36 questionnaire can provide a direct quantitative indication of an individual's health status and, as it is easy to administer, it has become the most wide-ly-used QOL evaluation tool in the world²⁻⁴. The SF-36 questionnaire was listed as an evaluation tool by an

international quality evaluation project in 1991⁵. Since then, the reliability and validity of the SF-36 questionnaire have been evaluated in a number of specific population world-wide⁶.

Quality of life (QOL) is either defined as the subjective perception of one's own well-being within socio-cultural context or as the satisfaction of desires and pleasures and the accomplishment of the ideal to a standard of perfection⁷. Great attention has been focused on the assessment of different populations ever since the concept of quality of life has become widely accepted by society⁸. However, the SF-36 questionnaire has seldom been used to assess the quality of life of Chinese medical students.

Most medical colleges follow a traditional Flexnerian curriculum in China. Medical training lasts 5 years and is divided into 2 years of basic sciences, 2 years of clinical medicine, and 1 year of internship in the hospital. Studies have reported that medical education and training have a negative impact on students' physical and mental health. As compared to the general population, medical students are more susceptible to depression⁹, anxiety¹⁰, stress¹¹ and burning out^{12, 13}. There may be a few influencing factors for this situation, such as academic courses and training14,15, curriculum which includes contact with patients, diseases and death^{16,17}. These negative effects will affect academic performance and interpersonal relationships, as well as the performance of medical practice in the future¹⁸. Dyrbye et al thought that student distress included four factors: anxiety, depression, stress and burning out. They put forward the theory that personal factors and medical school training factors have an impact on student distress¹⁹.

With the increasing enrollment of universities in China since 1999, the enrollment number of medical students, including junior college students of medicine, has increased from 75188 in 1998 to 533618 in 2010. The nearly 7-fold increase results in a serious shortage of teaching facilities, such as inadequate teachers, laboratories, classrooms and teaching hospitals^{20,21}. In addition, students do not have access to adequate logistical services, such as dormitories and canteens. Besides having to face those problems, Chinese medical students also face greater pressure regarding post-graduation employment²². Most graduates prefer not to work in the community or rural hospitals for lower salaries and poorer working conditions compared with those of larger hospitals. The large hospitals cannot provide enough jobs, so graduates face great employment pressure. Medical education is always long in duration and consists of great academic pressure and narrow professional employment opportunities²³. Some medical students with poor academic and professional performance are demoted or failed for the above issues. It has been speculated that these problems impair students' mental and physical health.

The quality of life (QOL) of Chinese medical students is of growing concern to educators and administrators in recent years²⁴. Assessing the quality of life of medical students can inform us of their health perspectives and health conditions and the related factors which have a impact on their quality of life (QOL). We can also take measures to increase their spare time and learning interests to improve their health conditions by medical education reforms, including improving curriculum and teaching methods. It has been reported that an integrated curriculum such as PBL may be less stressful for students²⁵.

The study was conducted among medical students in China Medical University in order to test the reliability and validity of the SF-36 questionnaire for measuring the QOL of Chinese medical students.

Materials and methods

Study Sample

In total, 1358 3rd year and 4th year medical students from 46 classes at 7 different areas of professional study were surveyed at China Medical University, using the cluster sampling method. These included 12 classes from clinical medicine, 2 classes from nursing, 2 classes from preventive medicine, 2 classes from forensic science, 2 classes from pharmacy, 2 classes from dentistry, and 1 classes from medical information management for each of the 3rd and 4th academic years. 1286 valid questionnaires were collected. The sample comprised of 599 males and 687 females. In the third and fourth years, students begin to study clinical courses, including surgery, pediatrics, obstetrics, and other disciplines. Although they have not gone into regular clinical rotations, they have begun to have some contact with patients during their clerkship courses in the hospitals.

Data Collection

This cross-sectional study was conducted on December 9, 2011. SF-36 questionnaires were distributed to students simultaneously at the end of their first period class. The surveyors took 5 minutes to explain the purpose of the study and the students were given 10 minutes to complete the questionnaire independently. The study was anonymous, and the results remain confidential. The completed questionnaires did not contain any identifying information about the individual subjects. Participation in the study was totally voluntary, and participants had the option of declining to answer specific questions or to leave the entire questionnaire blank if they did not wish to participate. The protocol was approved by the Bioethics Advisory Commission of China Medical University. All data were kept confidential and data protection was observed at all stages of the study.

Questionnaires

The questionnaire included socio-demographic information (gender, age, nationality, grade level, and specialty). If the students not understand any items,

the surveyors assisted them individually and recorded the relevant items for improvement if the questionnaire was used again. The SF-36 questionnaire includes 36 questions related to an individual's QOL that are summarized in two component summary scores, the Physical Component Summary (PCS) and the Mental Component Summary (MCS) scores.

Statistical Analysis

Validity was analyzed through collective validity, divisional validity and structural validity²⁵. After deducting the overlap between each of the 36 items and its related domain, the collective validity was considered to be good if the correlation coefficient remains > 0.4. To support the divisional validity, the items should have higher correlation with their hypothesized domains than with domains measuring other concepts. The statistical significance of the difference between the item-hypothesized domain and item-competing domain correlations was tested by the Steiger's t-test for dependent correlation²⁶. Factor analysis is a statistical method used to test the structural validity of a scale and describes variability among observed variables in terms of fewer unobserved variables - called factors²⁷. In the present study, factor analysis for the eight domains was used to evaluate the structural validity of the SF-36 questionnaire by testing whether the observed data for the eight domains collected during the study correlated with the hypothetical structure of the two overall component scores -- the PCS and the MCS. The Kaiser-Meyer-Olkin-Kriterium (KMO) statistic and Bartlett's spherical check were carried out to check for sample suitability for the factor analysis.

Internal reliability of the SF-36 questionnaire was measured by determining the internal uniformity, which is expressed by Cronbach's a coefficient. Cronbach's a coefficient was calculated for the eight domains of the SF-36 questionnaire and the reliability was considered to be adequate if the α value was >0.7²⁸.Split-half reliability, a measure of consistency where a test is split in two and the scores for each half of the test are compared with one another, was used to check the internal stability of the questionnaire, and test-retest reliability was used to assess the consistency of the questionnaire from one time to another^{29, 30}. In order to determine the test-retest reliability, a second round of evaluation was undertaken among 50 study subjects who were randomly selected 1 week later. The final analysis database was formed after analytical treatment for the logical error had been performed and abnormal values of the data had been obtained. The data were analyzed using SPSS® version 17.0 (SPSS Inc., Chicago, IL, USA) for Windows®.

A *P*-value of < 0.05 was considered to be statistically significant.

Results

Characteristics of Chinese Medical Students

Of the 1358 distributed questionnaires, 1286 (94.69%) were completed. The average age of the sample was 22.3 years (SD = 2.9), with 599 (46.5%) male students and 687 (53.5%) female students. The sample includes 654 3^{rd} year students and 632 4^{th} year students. There are 665 students from clinical medicine, 101 students from nursing, 114 students from preventive medicine, 116 students from forensic science, 116 students from pharmacy, 117 students from dentistry, and 57 students from medical information management.

Validity Analysis

Structural validity was evaluated by means of factor analysis, according to the degree of similarity between the hypothetical structure of the questionnaire conceived by researchers and the actual observed data. Results showed the KMO measure to be 0.786 and the Bartlett's spherical check to be χ^2 = 198.11 and P = 0.000, which taken together, indicated that the samples in this study were suitable for factor analysis. Factor analysis results indicated that when two component summary scores, the PCS and the MCS, were extracted from those of the eight domains whose characteristic roots were > 1 or approaching 1, the accumulative contribution rate was up to 73.6%. As shown in Table 1, the PCS has larger factor loads on PF, RP and BP domains with high correlation, and lower factor loads on RE and MH domains with low correlation in accordance with the theoretical hypothesis. The MCS has larger factor loads on VT, RE and MH domains with high correlation, but the social function domain where the factor loads for the MCS for the observed data were slightly low and appeared as only a moderate correlation that was not identical with the hypothesis. The correlation coefficient (r >0.50) for each item and its related domain, obtained by the correlation coefficient model, was relatively high, indicating good structural validity (Table 1).

As can be seen from Table 2, the coefficient range of the collective validity for all the domains except for SF and RE domains was >0.4. The collective validity and divisional validity were considered to be accepted.

Reliability Analysis

The degree of internal uniformity among the items, namely the correlation between the items and

the eight related domains, was expressed by Cronbach's α coefficient (Table 3). The overall Cronbach's α coefficient of the SF-36 questionnaire was 0.791, while the respective Cronbach's α coefficients for seven of the eight dimensions were > 0.70, excluding the social function dimension, which was 0.631. This met the requirement for group comparison. There was also a positive correlation between each of the eight domains of the SF-36 questionnaire (P < 0.01; Table 3). Table 3 shows the correlation coefficients (r) between the 36 items of the SF-36 questionnaire and the eight domains of study.

The retest of the correlation between the items showed that r > 0.70 could be achieved for all eight

domains (P < 0.01) (Table 4), demonstrating relatively good stability for the SF-36 questionnaire. The difference between the mean values for each domain after two rounds of measurements was not statistically significant.

The split-half reliability measure was determined by splitting the SF-36 items in each dimension by an odd-even split, calculating the correlation coefficient r_1 for each split separately and comparing the two, thereby calculating the reliability of each part of the split questionnaire. This was corrected using the Spearman-Brown prediction formula $r = 2r_1/(1 + r_1)$, which generated the value of r = 0.778 (P < 0.001), showing that this questionnaire was relatively stable.

Table 1. Structural validity of the SF-36: comparison of hypothetical correlation and factor analysis from data collected from Chinese medical students.

Domains	Hypothetical factor load		Practice factor load		
	PCS	MCS	PCS	MCS	
PF	+	-	0.762	0.228	
RP	+	-	0.786	0.186	
BP	+	-	0.731	0.283	
GH	*	*	0.678	0.287	
VT	*	*	0.496	0.762	
SF	*	+	0.621	0.612	
RE	-	+	0.315	0.713	
MH	-	+	0.241	0.859	

Correlation coefficient (*r*): $+ r \ge 0.70$; + 0.70 > r > 0.30; $-r \le 0.30$.

PCS, physical component summary; MCS, mental component summary; PF, physical function; RP, role limitations due to physical problems; BP, bodily pain; GH, general health; VT, vitality; SF, social function; RE, role limitations due to emotional problems; MH, mental health.

Table 2. The collective validity and divisional validity of each domain on the SF-36 questionnaire.

		Coefficient range		Collective validity		Divisional validity	
Domain	Number	Collective validity	Divisional validity	Successful number	Successful rate (%)	Successful number	Successful rate (%)
PF	10	0.445-0.657	0.095-0.412	10/10	100	80/80	100
RP	4	0.579-0.756	0.142-0.387	4/4	100	32/32	100
BP	2	0.754-0.897	0.121-0345	2/2	100	16/16	100
GH	5	0.579-0.704	0.085-0.478	5/5	100	40/40	100
VT	4	0.527-0.799	0.154-0.568	4/4	100	32/32	100
SF	2	0.279-0.461	0.128-0.367	1/2	50	16/16	100
RE	3	0.124-0.824	0.127-0.389	2/3	67	18/24	75
MH	5	0.487-0.789	0.175-0.489	5/5	100	40/40	100

PF, physical function; RP, role limitations due to physical problems; BP, bodily pain; GH, general health; VT, vitality; SF, social function; RE, role limitations due to emotional problems; MH, mental health.

	Internal uniform reliability of SF-36 items with domains		Positive correlation between SF-36 domains							
Domain	Test-retest reliability	Cronbach's <i>a</i> coefficient <i>n</i> = 1286	Domains							
n = 50	n = 50		PF	RP	BP	GH	VT	SF	RE	
PF	0.763*	0.826								
RP	0.727	0.745	0.268*							
BP	0.726*	0.793	0.302*	0.323*						
GH	0.701*	0.791	0.182*	0.271*	0.410*					
VT	0.862*	0.827	0.187*	0.197*	0.266*	0.322*				
SF	0.591	0.631	0.198*	0.282*	0.403*	0.301*	0.392*			
RE	0.701	0.732	0.274*	0.378*	0.274*	0.384*	0.370*	0.330*		
MH	0.738*	0.798	0.287*	0.184*	0.236*	0.191*	0.460*	0.321*	0.320*	

Table 3. The internal uniform reliability and correlation coefficient of the eight dimensions of the SF-36 questionnaire for measuring quality of life of Chinese medical students.

**P* < 0.01. PF, physical function; RP, role limitations due to physical problems; BP, bodily pain; GH, general health; VT, vitality; SF, social function; RE, role limitations due to emotional problems; MH, mental health.

Dimensions	Correlation coefficient (<i>r</i>)	<i>P</i> -value
PF	0.683 - 0.841	< 0.01
RP	0.712 - 0.811	< 0.01
BP	0.761 - 0.866	< 0.01
GH	0.712 - 0.842	< 0.01
VT	0.781 - 0.862	< 0.01
SF	0.598 - 0.643	< 0.01
RE	0.684 - 0.861	< 0.01
MH	0.695 - 0.813	< 0.01

Table 4. The correlation coefficient (r) between individual items and the eight domains of the SF-36 questionnaire.

PF, physical function; RP, role limitations due to physical problems; BP, bodily pain; GH, general health; VT, vitality; SF, social function; RE, role limitations due to emotional problems; MH, mental health.

Discussion

The SF-36 is considered to be a valid, reliable, concise, and generic measure of state of health that is potentially useful for application to students^{31, 32}. The present study was designed to test the reliability and validity of the SF-36 questionnaire in the population of Chinese medical students. Few studies have used validated instruments that reflect the multi-dimensional and subjective concept of QOL by eliciting information directly from students.

Compared with other questionnaires designed to evaluate QOL, the SF-36 questionnaire is short and flexible, which makes it much easier to administer. The SF-36 questionnaire can be completed manually or with the aid of a computer, by individuals, via a face-to-face interview or by telephone call with trained surveyors. The SF-36 questionnaire is widely used to monitor general population health status, to evaluate the efficacy of interventions, to monitor health status in patients with chronic disease and to determine the relative burdens of various diseases³³, ³⁴. Several health surveys have been evaluated in particular diseases and in special populations, such as the elderly, including populations in China^{35, 36}.

The results obtained indicate that the SF-36 is a useful measure and had good reliability and validity in the determination of the QOL among medical students, with relatively stable function. The Cronbach's a coefficients for all seven domains of the SF-36 was > 0.7, except where the coefficient for the social function dimension was 0.631. The items of the social function dimension may not be sensitive to cultural variations and may need to be modified according to the characteristics of the Chinese population. Previous study has also provided the evidence for this issue³⁷. In all, that indicates a good internal uniformity. The structural validity was in accordance with hypothetical correlations, which indicated a good overall structural validity.

In summary, we believe that this version of the SF-36 is suitable for its intended purpose to measure health in medical students, although some items that are related to social functions require more revision.

Further research is needed to identify the determinants of the QOL in medical students and improve their QOL in medical universities.

Conclusion

This study has provided psychometric properties of the SF- 36 health survey questionnaire and it can be used to assess the quality of life of Chinese medical students.

Abbreviations

QOL: Quality of life; SF-36: The 36-Item Short Form Health Survey.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant Number 71103200 and 30700690) and the Foundation of the Educational Department of Liaoning Province of China (Grant Number L2010577). The authors would like to express their gratitude to the teachers and students from student service office and students union to participate in related field work of this research.

Conflicts of interest

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

References

- Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30: 473-483.
- Lahmek P, Berlin I, Michel L, et al. Determinants of improvement in quality of life of alcohol-dependent patients during an inpatient withdrawal programme. Int J Med Sci. 2009 May 18; 6(4): 160-167.
- Meissner A, Stifoudi I, Weismüller P, et al. Sustained high quality of life in a 5-year long term follow-up after successful ablation for supra-ventricular tachycardia. results from a large retrospective patient cohort. Int J Med Sci. 2009; 6(1): 28-36.
- Kelm J, Bohrer P, Schmitt E, et al. Treatment of proximal femur infections with antibiotic-loaded cement spacers. Int J Med Sci. 2009 Sep 3; 6(5): 258-264.
- Brazier J. The SF-36 health survey questionnaire--a tool for economists. Health Econ. 1993; 2: 213-215.
- Corcoran WE, Durham CF. Quality of life as an outcome-based evaluation of coronary artery bypass graft critical paths using the sf-36. Qual Manag Health Care. 2000; 8: 72-81.
- The world health organization quality of life assessment (WHOQOL): Position paper from the world health organization. Soc Sci Med. 1995; 41: 1403-1409.
- Valenti M, Porzio G, Aielli F, et al. Physical exercise and quality of life in breast cancer survivors. Int J Med Sci. 2008 Jan 15; 5(1): 24-28.
- Dyrbye LN, Thomas MR, Eacker A, et al. Race, Ethnicity, and Medical Student Well-being in the United States. Arch Intern Med. 2007; 167: 2103-2109.
- Aktekin M, Karaman T, Senol YY, et al. Anxiety, depression and stressful life events among medical students: a prospective study in Antalya, Turkey. Med Educ. 2001; 35: 12-17.
- Compton MT, Carrera J, Frank E. Stress and depressive symptoms/dysphoria among US medical students: results from a large, nationally representative survey. J Nerv Ment Dis. 2008; 196: 891-897.
- Dyrbye LN, Thomas MR, Power DV, et al. Burnout and serious thoughts of dropping out of medical school: a multi-institutional study. Acad Med. 2010; 85: 94-102.

- Dyrbye LN, Thomas MR, Huschka MM, et al. A multicenter study of burnout, depression, and quality of life in minority and nonminority US medical students. Mayo Clin Proc. 2006; 81: 1435-1442.
- Paro HB, Morales NM, Silva CH, et al. Health-related quality of life of medical students. Med Educ. 2010; 44: 227-235.
- Moffat KJ, McConnachie A, Ross S, et al. First year medical student stress and coping in a problem-based learning medical curriculum. Med Educ. 2004; 38: 482-491.
- MacLeod RD, Parkin C, Pullon S, et al. Early clinical exposure to people who are dying: learning to care at the end of life. Med Educ. 2003; 37: 51-58.
- 17. Wear D. "Face-to-face with it": medical students' narratives about their end-of-life education. Acad Med. 2002; 77: 271-277.
- Dahlin M, Joneborg N, Runeson B. Stress and depression among medical students: a cross-sectional study. Med Educ. 2005; 39: 594-604.
- Dyrbye LN, Thomas MR, Shanafelt TD. Medical student distress: causes, consequences, and proposed solutions. Mayo Clin Proc. 2005; 80: 1613-1622.
- Department of Development and Planning of the ministry of education of China. Educational statistics yearbook of China. Beijing, China: People's Education Press; 1998.
- Department of Development and Planning of the ministry of education of China. Educational statistics yearbook of China. Beijing, China: People's Education Press; 2010.
- Liu KR, Hu GF, Zhang MY, et al. Psychological anxiety evaluation and analysis of graduates at a medical university under employment pressure. Nan Fang Yi Ke Da Xue Xue Bao. 2009; 29: 1071-1072.
- Wang XY, Rodriguez AC, Shu MR. Challenges to implementation of medical residency programs in China: a five-year study of attrition from west China hospital. Acad Med. 2010; 85: 1203-1208.
- Sun L, Sun LN, Sun YH, et al. Correlations between psychological symptoms and social relationships among medical undergraduates in Anhui Province of China. Int J Psychiatry Med. 2011; 42(1): 29-47.
- Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull. 1959; 56(2):81-105.
- Steiger JH. Tests for comparing elements of a correlation matrix. Psychol Bull. 1980; 87: 245-51.
- Ware JE Jr, Kosinski M, Gandek B, et al. The factor structure of the SF-36 heaith survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998; 51(11): 1159-1165.
- Fleiss JL. Design and Analysis of Clinical Experiments. New York, USA: John Wiley & Sons; 1986.
- Jafari H, Lahsaeizadeh S, Jafari P, et al. Quality of life in thalassemia major: reliability and validity of the Persian version of the SF-36 questionnaire. J Postgrad Med. 2008; 54: 273-275.
- Linde L, Sorensen J, Ostergaard M, et al. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D [corrected] RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol. 2008; 35: 1528-1537.
- Tyssen R, Vaglum P, Gronvold NT, et al. Suicidal ideation among medical students and young physicians: a nationwide and prospective study of prevalence and predictors. J Affect Disord. 2001; 64: 69-79.
- Givens JL, Tjia J. Depressed medical students' use of mental health services and barriers to use. Acad Med. 2002; 77: 918-921.
- Vilagut G, Valderas JM, Ferrer M, et al. Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components. Med Clin (Barc). 2008; 130: 726-735.
- 34. Onishi T, Nishikawa K, Hasegawa Y, et al. Assessment of health-related quality of life after radiofrequency ablation or laparoscopic surgery for small renal cell carcinoma: a prospective study with medical outcomes study 36-item health survey (SF-36). Jpn J Clin Oncol. 2007; 37: 750-754.
- Mok WY, Lam CL, Lam B, et al. A Chinese version of the sleep apnea quality of life index was evaluated for reliability, validity, and responsiveness. J Clin Epidemiol. 2004; 57: 470-478.
- Arslantas D, Ayranci U, Unsal A, et al. Prevalence of hypertension among individuals aged 50 years and over and its impact on health related quality of life in a semi-rural area of western Turkey. Chin Med J (Engl). 2008; 121: 1524-1531.
- Lam CL, Gandek B, Ren XS, et al. Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. J Clin Epidemioi. 1998 Nov; 51(11):1139-1147.