Reliability and Validity of the Turkish Version of SF-12 Health Survey

Cem SOYLU[1] , Bahtım KÜTÜK[2]

SUMMARY

Objective: The aim of this study is to determine the psychometric properties of the Turkish version of the SF-12 Health - a widely used scale that assesses health - related quality of life - and to assess whether its psychometric properties are as good as with the 36 items long version (SF-36).

Method: The study included a total of 190 participants including 138 young adults aged from 18 to 29 years (22.04±2.35) and 52 older adults aged from 60 to 80 years (67.27±5.22). The relationships between EQ-5D-3L, CES-D and SF-12 scales were investigated for content, convergent and divergent, criterion-related and the known-groups construct validity. In order to assess reliability, internal consistency and item-total correlation coefficients were calculated and test re-test analyses were conducted.

Results: The internal consistency coefficients of the scale were found to be satisfactory (for the components α=0.73 and 0.72). Consistent with the original study, physical and mental components of the SF-12-TR were found to be strongly correlated with the components of the SF-36 and explained, respectively, 86.6% and 92.4% of the variance. The components of the SF-12-TR were found to be moderately or strongly correlated to the sub-dimensions of the EQ-5D-3L and CES-D, which measure the same structures, while they were unrelated or weakly correlated with the sub-dimensions measuring different structures.

Conclusion: The findings indicated that the psychometric properties of the Turkish form of SF-12-TR are similar to the original version, and the SF-12-TR can be used as an alternative to the SF-36-TR in clinical and research settings in our country.

Keywords: The SF-12-TR, reliability, validity, the SF-36-TR

INTRODUCTION

Health-related quality of life is a multi-dimensional concept comprisingthe physical, psychological and social aspects of health and it is generally accepted as an important measure of the outcomes of health services (Globe et al. 2002). Measurement toolsdeveloped to assess health-related qualityof life are divided into two main groups as the disease-specific and the generic tools (Coons et al. 2000, Michalos 2014). The disease-specific tools are typically designed only to assess the effect of a particular disease or treatment on quality of life and, therefore, they do not provide comparative data between different disease groups or between disease groups and the general population. In contrast, the generic assessment tools, not being dependent on a specific disease or treatment, can be used to evaluate the quality of life of both patients and healthy individuals on a wide area of application including primarily health screening, research on disease burden and epidemiology and studies conducted to define health status and develop individual treatment plans (Coons et al. 2000, Wells et al. 2011).

In the literature, 'SF-36 Health Survey', 'Nottingham Health Profile', 'European Quality of Life Five Dimension Questionnaire' (EuroQol EQ-5D-3L) and 'Sickness Impact Profile' are among the most frequently used scales to evaluate health-related quality of life in both the patient group (Cooper et al. 2017) and non-clinical, in other words, the general population (Coons et al. 2000). Among these, the SF-36 was developed in 1993 by the Medical Outcome Study (MOS) team including researchers from different disciplines to evaluate the functional status and well-being of individuals (Ware et al. 1993).

An important characteristic of the SF-36, emphasised by researchers, is its applicability to provide a standard value on comparative research on clinical and non-clinical groups with different characteristics. Hence, the SF-36 has been one of the most

widely used instruments to measure health-related quality of life both in normal population and in various patient groups since its development (Hagell and Westergren 2011, Huang et al. 2013). The SF-36 and the SF-12, its shortened form, were identified in a recent systematic review article as the most frequently used generic measurement toolson patient groups (Reynaud et al. 2020). Another systematic review of 63 studies using generic measurement tools to assess quality of life among adults in the general population pointed out that 33.3% of the research had used the SF-36 (Pequeno et al. 2020).

The SF-36 comprises a total of 36 items constituting the eight sub-dimensions, namely, 'physical and social functioning', 'role-physical and role-emotional', 'general and mental health', 'bodily pain and energy'(Ware et al. 1993). Theeight sub-dimensions can be scored separately or thescores can be combined intophysical health component summary scores (PCS) and mental health component summary scores (MCS) to provide a general evaluation(Ware et al. 1993, Ware et al. 1994). Despite its versatility and frequent use, theSF-36 has been criticized for its high number of items and long duration of application (Globe, 2002). With the intentto develop a practical measure applicable in shorter time, Ware et al. (1995) developed the SF-12, which contains the same dimensions as the SF-36 but has less questions, thus providing a significant advantage in its usage. The randomized controlled study by Kim et al. (2017) asking patients to complete both the SF-36 and the SF-12 showed that the ratio of items completed in the two forms was significantly higher in the SF-12. Also, the evidence provided by comparative research that both the SF-36 and the SF-12 had similar psychometric characteristics providing close results, contributed to the popularity of the SF-12 (Lam et al. 2005).

Indeed, SF-12 has been translated into different languages such as Russian (Hoffmann et al. 2005), Greek (Kontodimopoulos et al. 2007), and Chinese (Lam et al. 2005) and proven to be a valid and reliable tool for use in the general population (Younsi and Chakroun 2014), on refugees (Hoffmann et al. 2005), various patient groups (Globe 2002, Hagell and Westergren 2011, Huang et al. 2013, Silverberg et al. 2019) and adults in different developmental stages of life, especially among the young and the elderly adults (Schofield and Mishra 1998, Shou et al. 2016), thus becoming one of the most important measurement tool world wide. Recent evaluation of the psychometric properties of the SF-12 in Nigeria on patientswith chronic low back pain showed satisfactory level of known-group, convergent and divergent validity and reliability on the basis of test-retest and internal consistency coefficients (Ibrahim et al. 2020). A similar study in Kenya has proven that SF-12 can be used to assess the physical and mental health of patients with HIV/AIDS (Patel et al. 2020). Younsi (2015) investigated the psychometric properties of the SF-12 scale in a study, which included approximately 4 thousand people aged 18 and older

and found that it is a reliable and valid scale for measuring health-related quality of life in the Tunisian population.

Adequate level of validity and reliability of the SF-36-TR has been demonstrated in Turkey, on patients (Koçyiğit et al. 1999) and on normal healthy groups (Demiral et al. 2006). However, a study on the validity and reliability of the SF-12 adapted in the Turkish language has not been available. The aim of this study is to examine the psychometric properties of the SF-12-TRbased on the SF-36-TRand to compare the results obtained with both forms in order to demonstrate that SF-12-TR can be used as an alternative to the SF-36-TR.

METHOD

Participants

Being literate and an adult of 18-30 and ≥ 60 years of age, as the young adult and the older adultparticpants, respectively, constituted the inclusion criteria of the study. During April-July 2019, a total of 208 eligible participants were reached, of whom 5 were excluded on grounds of being treated for dementia or schizophrenia; 6 were excluded for not completing the sociodemographic questionnaire or other test scales; and 7 were excluded for entering outlier data. The dataof the study wereacquired from 190 adults including 138 young adults aged between 18 and 29 (22.04±2.35), and 52 community-dwelling older adults aged between 60 and 80 (67.27±5.22). The study focused on both younger and older adults as significant differences were expected in the overall health status between these development stages of life. The socio-demographic characteristics of the two groups of participants are shown in Table 1. Only 3 (5.8%) of a total of 52 older adult participants were employed, 24 (46.2%) were unemployed, and 25 (48.1%) had retired.

Data Acquisiton Tools

The Socodemographic questionnaire (SDQ)

The SDQ was designed for this research to obtain personal information on the age, gender, marital status and diagnosed chronic diseases of the participants.

The SF-12-TR

Similarly to the SF-36-TR, theSF-12-TR consists of a total of 12 items in 8 subdimesions on physical functioning (2 items), role-physical (2 items), bodily pain (1 item), general health (1 item), energy (1 item), social functioning (1 item), role-emotional (2 items), and mental health (2 items). The items on role-physical and role-emotional are dichotomous, to be answered as "yes or no", while theother items are Likert-type with response options ranging from 3 to 6. The total PCS-12 score is obtained from the subdimensionson general health, physical functioning, role-physical, and

Table 1. Socio-demographic Characteristics of the Participants Grouped by Age Groups

Variables	Younger adults		Older adults		χ^2	df	P
	n	%	n	%			
Gender					4.28	1	0.039
Female	94	68.1	27	51.9			
Male	44	31.9	25	48.1			
Marital status					27.55	1	<0.001
Single or without a partner	80	58	8	15.4			
Married or with a partner	58	42	44	84.6			
Number of chronic disease					116.12	2	<0.001
No disease	103	74.6	0	0			
1 disease	34	24.6	23	44.2			
2 or more disease	1	0.7	29	55.8			
Education					-	-	-
Literate	-	-	4	7.7			
Primary Sch.	-	-	28	53.8			
Secondary Sch.	-	-	6	11.5			
Lyceé	138	100	9	17.3			
University	-	-	5	9.6			
	Mean±SD		Mean±SD		t	df	p
Chronological age	22.04±2.35		67.27±5.22		-60.18	58.95	<0.001

bodily pain; and the total MCS-12 score is obtained from the subdimensions on social functioning, role-emotional, mental health and energy. Both the PCS-12 and MCS-12 scores range from 0 to 100, with higher scores indicating better health status.

The developers of the original version of the SF-36, having observed that the norm values used for scoring yielded negligible differences between the results of participants from 9 different countries tested together, concluded that the USA norms with a mean of 50 and SD of 10 should be used in order to obtain interculturally comparable and interpretable scores (Gandek et al. 1998).

Thus, the component summary scores of the SF-36 and SF-12 were obtained in accordance with the manuals developed for these tools, by using the USA norm values and standard algorithms(Ware et al. 1993, Ware et al. 1994, Ware et al. 1995). Although theSF-36 format allows assigning new values to replace missing data, in this study only the fully completed SF-12-TR forms were used in data anlysessince some subdimensions of the SF-12 are measured with only one item.

TheEuropean Quality of Life Five Dimension (EuroQol EQ-5D-3L)

This measurement tool, developed by the European Quality of Life Research Society (EuroQol) in 1990 to measure health-related quality of life, consists of the five dimensions on mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has one item with three score levelsanswered as *no problems, some problems and extreme problems*. There is also a visual analog scale (VAS) marked on a thermometer-like scale by giving a score between 0 and 100

about theparticipant's current health status. The validity and reliability of the SF-36-TR was determined by Kahyaoğlu Süt and Ensar (2011). It had valid and reliable properties with an internal consistency coefficient of 0.86 which was found to be 0.64 in the present study.

The Epidemiologic Studies Depression Scale (CES-D)

This scale, developed by Radloff (1977), consists of 20 items with 4 rating points ranging from 0 to 3. The 4th, 8th, 12th and 16th items of the scale are reverse items; and the obtainable total score varies between 0 and 60. A high score indicates the presence of depression. The scale was adapted to the Turkish language by Tatar and Saltukoğlu (2010) and the internal consistency coefficient was found to range between 0.75 and 0.90. The internal consistency coefficient was calculated to be 0.93 in the present study.

Procedure

John E. Ware and Barbara Gandek, whose permissions were sought as members of the team of developers of the SF-12, recommended not to retranslate the SF-12 items, as these were already embedded in the SF-36-TR, but to use the SF-36-TR and to baseon it the calculation of the results on the 12 items of the SF-12-TR, as was done by others (see Globe et al. 2002, Kontodimopoulos et al. 2007, Lam et al. 2005). These recommendations were followed in this study. The Ethics Board and Commission of Ege University approval numbered 04/12-206 and dated 28.03.2019 was obtained for the study. The older group of adultswas reached by convenience sampling, starting from the immediate surroundings of the researchers. The young adult group participants comprised the volunteering

undergraduate students of the Ege University Department of Psychologywho had been reached collectively. Those enrolled in the study were given the set of data acquisition tools and asked to fill them in the classroom environment. All participants of the study gave written informed consent for inclusion in the study and wereinformed that they could stop completing the data acquisition forms at any time.

Statistical Analyses

Beforestarting the statistical analyses, controls were undertaken to eliminate errors on data entries and that the variables were at the expected intervals. Normality assumptionsin the distribution of continuous variables were examined by Histogram and Normal Q-Q plotsand statistical techniques including the Kolmogorov Smirnov test, Skewness and Kurtosis values (Tabachnick and Fidell, 2013). The relationship between the two age groups and the categorical variables including gender, marital status and the number of chronic diseases wereanalysed using the 2X2 and 2X3 chi-square analyses. The reliability of the SF-12-TR was determined by calculating the Cronbach's alpha internal consistency coefficient, intraclasscorrelation coefficients, item-total correlation coefficients and the correlation coefficients between the PCS-12 – PCS-36 and the MCS-12 – MCS-36 scores.

The content validity of the scale was assessed by the ratio of total variance of the PCS-36 and MCS-36 explained by the SF-12-TR scores using simple linear regression analysis. The concurrent validity was tested by the one-way ANCOVA. The known-groups validity and the criterion validity of the SF-12-TRwereanalysed by the Independent Samples t-test and

Dependent Samples t-test. The convergent and divergent/discriminant validity analyses were made using the Pearson correlation analysis. The IBM SPSS.23 program and R (Version 3.3.2 R Core Team, 2016) was used for the statistical analyses and the $p<0.05$ valuewastaken to indicate statistical significance.

RESULTS

Reliability of the SF-12-TR

The internal consistency coefficient calculated to evaluate the reliability of theSF-12-TR was found to be 0.73 for the PCS-12 and 0.72 for the MCS-12. In order to determine the consistency of the scores obtained from the scale over time, testing was repeated after an interval of 2 weeks with 62 young adult participants and the intraclass correlation coefficients for the test-retest reliability were found to be 0.73 for the PCS-12 and 0.72 for the MCS-12.

The reliability of the SF-12-TR was also evaluated by calculating the item-total score correlation coefficients for the PCS-12 and MCS-12, and the correlation coefficients were calculated between the PCS-12 – PCS-36 and the MCS-12 – MCS-36. As expected, item-total score correlation coefficients for the PCS-12 ranged between0.40 and 0.55 and item-total score correlation coefficients for the MCS-12 ranged between 0.32 and 0.67. A statistically significant strong correlationwas determined between the PCS-12 and the PCS-36 (r=0.93; p<0.001) and between the MCS-12 and the MCS-36 (r=0.96; p<0.001). These results have indicatedacceptable reliability coefficients for the SF-12-TR.

Table 2. Results of Simple Linear Regression Analysis for Prediction of the FCS-36 and MCS-36 by the FCS-12 and MCS-12

Variables	B	SE	t	P	%95 Confidence Interval	R²	ΔR²
PCS-36							
Constant	-1.85	1.55	-1.20	0.233	[-4.90 – 1.20]		
PCS-12	1.04	0.03	34.99	<0.001	[0.98 – 1.10]	0.867	0.866
MCS-36							
Constant	-1.87	0.90	-2.09	.038	[-3.63 – -0.11]		
MCS-12	1.02	0.02	47.89	<0.001	[0.98 – 1.06]	0.924	0.924

Table 3. Correlation Coefficients of PCS-12 and MCS-12 with the CES-D and the Sub-dimensions of the EQ-5D-3L

Variables	PCS-12	MCS-12	EQ-5D-3L Mobility	EQ-5D-3L Self-care	EQ-5D-3L Usual activities	EQ-5D-3L Pain/discomfort	EQ-5D-3L Anxiety/depression	CES-D
PCS-12	-							
MCS-12	-0.13	-						
EQ-5D-3L Mobility	-0.50**	-0.07	-					
EQ-5D-3L Self-care	-0.36**	-0.12	0.50**	-				
EQ-5D-3L Usual activities	-0.36**	-0.45**	0.36**	0.34**	-			
EQ-5D-3L Pain/discomfort	-0.50**	-0.22**	0.44**	0.35**	0.38**	-		
EQ-5D-3L Anxiety/depression	0.06	-0.70**	0.02	0.05	0.39**	0.18*	-	
CES-D	-0.14	-0.48**	0.12	0.21**	0.41**	0.13	0.45**	-

*p<0.05 **p<0.01.

Validity of the SF-12-TR

The SF-12-TR Content validity: Was measured by calculating the ratio of the total variance explained by the PCS-12 and MCS-12 in the PCS-36 and MCS-36. Simple linear regression analyses showed that the PCS-12 significantly predicted the PCS-36 and explained 86.6% of the total variance $F(1.188)=1224.60$; $p<0.001$, $R^2=0.867$; $R^2_{Adjusted}=0.866$ and that the MCS-12 significantly predicted the MCS-36 and explained 92.4% of the total variance $F(1.188)=2293.49$; $p<0.001$, $R^2=0.924$, $R^2_{Adjusted}=0.924$ (Table 2).

The SF-12-TR Convergent and divergent/discriminant validity: Were determined by analyzing the relationships of the PCS-12 and MCS-12 with the scores on the EQ-5D-3L subdimensions and the CES-D. The results are shown in Table 3. The PCS-12 was moderately correlated with the mobility, self-care, usual activities, and pain/discomfort subdimensions of the EQ-5D-3L, supporting the convergent validity. The SF-12 divergent/discriminant validity was supported by determining that the PCS-12 did not correlate with the anxiety/depression subdimension score of the EQ-5D-3L and the CES-D scale. The MCS-12 was found to be strongly correlated with the EQ-5D-3L anxiety/depression subdimension and moderately correlated with the CES-D scale which supported the SF-12 convergent validity. The MCS-12 did not correlate with the EQ-5D-3L mobility and self-care subdimensions, supporting the divergent validity of the SF-12.

Also, the divergent/discriminant validity of the SF-12-TR was indicated by the expected lack of significant correlation between the PCS-12 and MCS-12.

The SF-12-TR Known-groups construct validity: Was assessed by the differences in the PCS-12 and MCS-12 scores according to the sociodemographic variables (Table 4). As

Table 4. Components of the SF-36-TR- and the SF-12-TR According to Socio-demographic Characteristics (Mean ±Standard Deviation)

Variables	Physical component summary (PCS) score		Statistic*	Mental component summary (MCS) score		Statistic*
	SF-36	SF-12		SF-36	SF-12	
Age group						
Younger	54.15±7.20	53.77±6.44	t(137)=1.31; p=0.194	36.67±11.71	38.37±11.12	t(137)=-6.07; p<0.001
Older	44.85±10.40	45.12±9.10	t(51)=-0.60; p=0.554	46.23±11.64	46.01±11.63	t(51)=0.47; p=0.640
Statistic**	t(70,25)=5.94; p<0.001	t(71,12)=6.29; p<0.001		t(188)=-5.03; p<0.001	t(188)=-4.17; p<0.001	
Gender						
Female	50.64±9.57	50.64±8.40	t(120)=-0.01; p=0.992	39.10±12.04	40.17±11.09	t(120)=-3.30; p=0.001
Male	53.29±8.21	52.73±7.74	t(68)=1.69; p=0.095	39.62±13.14	40.99±12.86	t(68)=-3.56; p=0.001
Statistic**	t(188)=-1.93; p=0.055	t(188)=-1.69; p=0.092		t(188)=-0.28; p=0.781	t(188)=-0.46; p=0.643	
Marital status**						
Single	52.94±8.13	52.43±7.40	t(87)=1.28; p=0.203	36.73±12.44	38.44±11.58	t(87)=-5.07; p<0.001
Married	50.45±9.87	50.51±8.78	t(101)=-0.21; p=0.831	41.49±12.03	42.21±11.65	t(101)=-2.02; p=0.046
Statistic**	t(187,61)=1.91; p=0.058	t(188)=1.62; p=0.108		t(188)=-2.68; p=0.008	t(188)=-2.23; p=0.027	
Chronic disease						
None	55.01±6.52	54.39±5.58	t(102)=1.99; p=0.049	36.32±11.81	38.06±11.33	t(102)=-5.28; p<0.001
1 disease	50.04±9.09	50.21±8.47	t(56)=-0.33; p=0.741	42.71±12.48	43.10±11.81	t(56)=-0.91; p=0.368
2 or more	42.85±10.65	43.38±9.47	t(29)=1.00; p=0.328	42.97±12.05	43.71±11.51	t(29)=-1.01; p=0.293
Statistic***	F(2,186)=18.96; p<0.001	F(2,186)=18.98; p<0.001		F(2,186)=4.45; p=0.013	F(2,186)=3.49; p=0.033	

* The p value obtained from the dependent samples t-test
** The p value obtained from the Independent samples t-test
*** The p value obtained from the One-way ANCOVA. Educational status has been changed as the duration of education and taken as the covariate variable.
**** The singles group includes the younger adult participants divorced, widowed, living separately or without a partner, the married group includes the younger adult participants living with a partner.

expected, the mean PCS-12of 53.77±6.44 obtained by the young adult group was significantly higher than the mean PCS-12 of 45.12±9.10 obtained by the older adult group [t(71.12)=6.29; p<0.001], whilethe mean MCS-12 of 38.37±11.12 obtained bythe young adults was significantly lower than the mean MCS-12 of 46.01±11.63 obtained by the older adults [t(188)=- 4.17; p<0.001]. Also, the mean MCS-12 of 42.21±11.65 obtained bythe married adults was significantly higher than the mean MCS-12 of 38.44±11.58 obtained by the single adults [t(70.25)=5.94; p<0.05]; while there was not asignificant difference between the MCS-12 valuesofthe married and single adults (p>0.05). There were not gender based differences inthe PCS-12 and the MCS-12 values of the participants of the young and the older adult participants (p>0.05).

The SF-12 Criterion validity: Was assessed on the basis ofwhether the PCS-12 and MCS-12 differed or were similar when compared to the PCS-36 and MCS-36using the dependent samples t-test. As shown in Table 4, similarities were observed in most of the analysesin thethelevel and direction of statistical significance or insignificance betweenthePCS-36 and MCS-36 and betweenthe PCS-12 and MCS-12 of the participants compared on the basis of age, gender and marital status.

The SF-12-TR Concurrent validity: The effect of the number of chronic diseases of the participants on the PCS-12 and MCS-12 and the presence of any difference between these component summary scores and those of the participants without chronic diseases were assessed by using the one-way ANCOVA. As expected, after controlling for the demographic factor of education duration, the number of chronic diseases were seen to have a statisticallysignificant effect on the PCS-12 [F(2.186)=18.98; p<0.001]. TheSidak-corrected Post-Hoc comparisons showed that the mean PCS-12 of the participants without chronic diseaseswas significantly higher than the PCS-12of the participants with 1 and ≥2 chronic diseases [(respectively, (54.39±5.58), (50.21±8.47) and (43.38±9.47)]. The number of chronic diseases also had astatistically significant effect on the MCS-12 after controlling for education duration[F(2.186)=3.49; p<0.05]. TheSidak-corrected Post-Hoc comparisons showedthat the mean MCS-12 of the participants without chronic diseases was significantly lower than those of the participants with 1 and ≥2 chronic diseases [(respectively, (38.06±11.33), (43.10±11.81) and (43.71±11.51)].

DISCUSSION

This study aimed to examine the psychometric properties of the Turkish language version of the SF-12 (SF-12-TR) which is one of the most frequently used scales to evaluate health-related quality of life.

Reliability Analyses

Whereas the internal consistency coefficient calculated on theMCS-12 was of a satisfactory level and in agreement with that obtained by the original study, the internal consistency coefficient calculated on the PCS-12, although satisfactory in being >0.70, was relatively lower than that given in the original studybyWare et al. (1995). The interval consistency coefficients calculated for the Russian language version the SF-12 were relatively lower as compared to the results of our study (Hoffmann et al. 2005).

Investigation of the validity and reliability of the SF-12 in Nigeria found the test-retest reliability coefficentsto be 0.79 and 0.85 for the PCS-12 and MCS-12, respectively (Ibrahim et al. 2020). These values are higher than 0.73and 0.72 for the PCS-12 and the MCS-12, respectively, obtained in our study. This could have been due to changes in the health status of the particpantswithin the 2-week interval taken for re-testing. Having obtained satisfactory internal consistency coefficients and also satifactorytest-retest reliability coefficients for the similar level of scores measured 2 weeks apartsupport the reliability of the SF-12-TR.

The item-total score correlation coefficients were reported to range between 0.54 and 0.67 during testing the validity and reliability of the SF-12 on elderly Chinese participants(Shou et al. 2016). Another study on 3 different groups consisting of 'elderly participants', 'Parkinson's disease patients' and 'stroke patients', reported the item-total score correlation coefficients to range between 0.38 and 0.77 (Jakobsson et al.2012). These values are similar to those obtained in our study. Given thati-tem-total score correlation coefficient values should be ≥0.30 (Büyüköztürk, 2017), the values obtained in our study for SF-12-TR are, therefore, acceptable.

The developers of the SF-12 reported a high level of correlationbetween the PCS-12 and PCS-36 (r=0.95) and between the MCS-12 and MCS-36 (r=0.97) (Gandek et al. 1998). A recent study reported similarlyhigh levels of correlation between the PCS-12 and the PCS-36 (r=0.88) and between the MCS-12 and the MCS-36 (r=0.93)of a clinical group of participants (Webster and Feller 2016). Significant correlations were also reported between the PCS-12 and the PCS-36 (r=0.94) and between the MCS-12 and the MCS-36 (r=0.96) of diabetes patients (Wukich et al. 2016). In agreement with these results of others, we also determined statistically significant correlationsbetween the PCS-12 and the PCS-36 (r=0.93) and between the MCS-12 and the MCS-36 (r=0.96) of the participants of our study, which confirmthe reliability of the SF-12-TR.

Validity Analyses

Analysis of the content validity of the Chinese language version of the SF-12showed that the PCS-12 and MCS-12

explained, respectively, 88% and 90% of the variance in the PCS-36 and MCS-36 (Lam et al. 2002). The corresponding figures explaining the variance, as obtained for the Greek language version of the SF-12, were, respectively, 93.2% and 86.9% (Kontodimopoulos et al. 2007). In agreement with these results of others, our results on the FCS-12 and MCS-12 explained 86.6% and 92.4% of the variance in the FCS-36 and MCS-36, respectively, and supported the content validity of the SF-12-TR.

Investigation of the correlations between the SF-12 and the EQ-5D-3L scores in pregnant females with HIV showed strong correlations between the MCS-12 and the scores on the anxiety/depression subdimension of the EQ-5D-3L, and between the PCS-12 and the functional subdimensions of the EQ-5D-3L(Wang et al. 2017). In agreement with these results, ourstudy demonstrated significant correlations betweenthe PCS-12 and scores on the mobility, self-care, usual activities, and pain/discomfort subdimensions of the EQ-5D-3Land between theMCS-12 and thescores on the CES-D and on the anxiety/depression subdimensionof the EQ-5D-3L, supporting the convergent validity of the SF-12-TR.

Analyses for Divergent validity showed that the PCS-12 did not significantly correlate with the scores on the CES-D and on the anxiety/depression subdimension of the EQ-5D-3; and that the MCS-12 also did not significantly correlate with scores on the mobility and self-care subdimensions of the EQ-5D-3L. Our results areconsistent with the reports of others in literature on the relationships between the SF-12 and the CES-D and the EQ-5D-3Lsubdimensions (Kontodimopouloset al. 2007, Lu and Liang 2016, Yu et al. 2015). A very low and nonsignificant relationship was demonstrated between the PCS-12 and MCS-12 (r=0.06; p>0.05) by a study on the psychometric properties of the SF12 (Ware et al. 1996). As expected, a very low and nonsignificant relationship was also demonstrated between the PCS-12 and MCS-12 (r=-0.13; p>0.05) in our study. Given this result and those on the correlation coefficients between the PCS-12 and MCS-12 and the scores on the CES-D and the EQ-5D-3L subdimensions, it can be said that SF-12-TR has convergent and divergent/discriminant validity.

Within the scope of developing the SF-12, investigation on 3 different age groups of 18-44, 45-64, and ≥65 years showed that the PCS-12 of the 18-44 age group had significantly higher PCS-12 and MCS-12 wassignificantly lower than those of the group with ≥65years of age (Ware et al. 1995). Our study comparing young adults of 18-30 years of age with the elderly participants of ≥65 years of age, also demonstrated that the younger adults had higher PCS-12 and lower MCS-12 as compared to the older adults. These results support the known-groups validity of the SF-12-TR, and indicate that both physical and mental health are significantly affected by age.

Investigation of the known-groups validity of SF-12 in members of the Tunisian population showed that the PCS-12 and MCS-12 were significantly higher in the male (51.9%) than the female (48.1%) participants; and in the married participans as compared to divorced and widowed participants (Younsi and Chakroun2014). Contrary to the expectations, our study including 36.3% male participantsdid not determine statistically significant gender based differences between the PCS-12 and MCS-12 of the participants. Also, married participants in our study had significantly higher MCS-12 than the single participants, while their PCS-12 values did not differ significantly. Although the results did not meet expectations, research has indicated that marital status affects mental rather than physical health(Wójcik et al. 2019).

The SF-36 was tested on patients with retinal diseases in order to evaluate the criterion validity of the SF-12 and the SF-12 and SF-36 composite summary scores were compared between four patient groups categorized on the type of the retinal disease. It was reported that the SF-36 and SF-12 component summary scores were generally similar to each other and gave similar significance values. For example, the PCS-36 and PCS-12 scores were 50±10 and 50±9, respectively, and the MCS-36 and MCS-12 scores were 53±10and 52±10, respectively, with p>0.5 for both sets of resultsof the posterior vitreous detachment patient group. These results were believed to support the criterion validity of the SF-12 (Globe et al. 2002). When the component summary scores of the Greek language version of the SF-36 were compared on the basis of demographic features including gender, age, education level and presence chronic disease with those of the SF-12 form derived from the SF-36 form, the results were generally very close to each other. For example, the PCS-36 and the PCS-12 of the female participants were, respectively, 48.74±12.06 and 48.44±10.75; and the MCS-36 and the MCS-12 scores were, respectively, 46.90±9.88and 47.98±9.84, supporting the criterion validity of the SF-12 (Kontodimopoulos et al. 2007). In agreement with these results, similar scores and significance levels were obtained in our study when the PCS-36 and MCS-36 and the PCS-12 and MCS-12 of the participants were compared on the basis of different sociodemographic features, thus supporting the criterion validity fo the SF-12-TR.

Limitations of the Study

Although individuals in two different developmental stages of lifewere included in the studywith the expectation to observe a significant difference in health status, the presence of asmall number of older adults as compared to the younger adult groupis considered to be a limitation of this study. Moreover, given the exclusion of middle aged adults with diverse educational backgrounds, having young adult participants with a homogenous educational background and the participant

groups not being representative of the general population by including only 36.3% males are among the limitations that should be amended in future studies by working on participant groups that better respresent the demographics of the general population.

CONCLUSION

When the resultsobtained in thisstudy are evaluated as a whole, it is seen thatthe SF-12-TRis a valuable toolwith adequate psychometric properties to measure general health status and health-related quality of life in clinical practice and scientific research in our country. It is thought that the contribution to our literature of an instrumentin the Turkish language enabling the measurement of health status in a short time with relatively few questions in both general population groups and groups with specific diseases will increase the number ofapplications in the field. Finally, as recommended by the developers ofthe SF-12, usingthe USA norm scores to address psychometric properties should also enable cross-cultural comparisons.

REFERENCES

Büyüköztürk Ş (2007) Sosyal Bilimler için Veri Analizi El Kitabı (8. Baskı). Ankara, Pegem Akademik Yayıncılık, s. 171.

Coons SJ, Rao S, Keininger DL et al (2000) A comparative review of generic quality-of-life instruments. Pharmacoeconomics 17:13-35.

Cooper V, Clatworthy J, Harding R et al (2017) Measuring quality of life among people living with HIV: a systematic review of reviews. Health Qual Life Outcomes 15:220.

Demiral Y, Ergör G, Ünal B et al (2006) Normative data and discriminative properties of short form 36 (SF-36) in Turkish urban population. BMC Public Health 6:247.

Gandek B, Ware JE, Aaronson NK et al (1998) Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. J ClinEpidemiol 51:1171-8.

Globe DR, Levin S, Chang TS et al (2002) Validity of the SF- 12 quality of life instrument in patients with retinal diseases. Ophthalmology 109:1793-8.

Hagell P, Westergren A (2011) Measurement properties of the SF-12 Health Survey in Parkinson's disease. J Parkinsons Dis 1:185-96.

Hoffmann C, McFarland BH, Kinzie JD et al (2005) Psychometric properties of a Russian version of the SF-12 Health Survey in a refugee population. Compr Psychiatry 46:390-7.

Huang Q, Zhou ZK, Shen B et al (2013) Validation of the SF-12 among adult Kashin-Beck disease patients in Aba Tibetan autonomous area in China. RheumatolInt 33:719-24.

Ibrahim AA, Akindele MO, Ganiyu SO et al (2020) The Hausa 12-item short-form health survey (SF-12): Translation, cross-cultural adaptation and validation in mixed urban and rural Nigerian populations with chronic low back pain. PloS One, 15:e0232223.

Jakobsson U, Westergren A, Lindskov S et al (2012) Construct validity of the SF-12 in three different samples. J Eval Clin Pract 18:560-6.

Kahyaoğlu Süt H, Ünsar S (2011) Is EQ-5D a valid quality of life instrument in patients with acute coronary syndrome? Anatol J Cardiol 11:156-62.

Kim PJ, Kumar A, Elmarsafi T et al (2017) Comparison of completion rates for SF-36 compared with SF-12 quality of life surveys at a Tertiary Urban Wound Center. J Foot Ankle Surg 56:1031-5.

Koçyiğit H, Aydemir Ö, Fişek G et al (1999) Kısa Form-36 (KF36)'nınTürkçe Versiyonunun Güvenilirliği ve Geçerliliği. İlaç ve Tedavi Dergisi 12:102-6.

Kontodimopoulos N, Pappa E, Niakas D et al (2007) Validity of SF-12 summary scores in a Greek general population. Health Qual Life Outcomes 5:55.

Lam CL, Eileen YY, Gandek B (2005) Is the standard SF-12 health survey valid and equivalent for a Chinese population? Qual Life Res 14:539-47.

Lu P, Liang Y (2016) Health-related quality of life of young Chinese civil servants working in local government: comparison of SF-12 and EQ5D. Appl Res Qual Life 11:1445-64.

Michalos AC (2014) Encyclopedia of quality of life and well-being research. Dordrecht: Springer Netherlands, s. 265.

Patel AR, Lester RT, Marra CA et al (2017) The validity of the SF-12 and SF-6D instruments in people living with HIV/AIDS in Kenya. Health Qual Life Outcomes 15:143.

Pequeno NPF, Cabral NLA, Marchioni DM et al (2020) Quality of life assessment instruments for adults: a systematic review of population-based studies. Health Qual Life Outcomes 18:208.

R Core Team (2016) R: A language and environment for statistical computing. Austria, R Foundation for Statistical Computing.

Radloff LS (1977) The CES-D scale: A self-report depression scale for research in the general population. Appl Psych Meas 1:385-401.

Reynaud V, Verdilos A, Pereira B et al (2020) Core Outcome Measurement Instruments for Clinical Trials of Total Knee Arthroplasty: A Systematic Review. J Clin Med 9:2439.

Schofield MJ, Mishra G (1998) Validity of the SF-12 compared with the SF-36 Health Survey in pilot studies of the Australian Longitudinal Study on Women's Health. J Health Psychol 3:259-71.

Shou J, Ren L, Wang H et al (2016) Reliability and validity of 12-item Short-Form health survey (SF-12) for the health status of Chinese community elderly population in Xujiahui district of Shanghai. Aging ClinExp Res 28:339-46.

Silverberg JI, Gelfand JM, Margolis DJ et al (2019) Validation and interpretation of Short Form 12 and comparison with Dermatology Life Quality Index in atopic dermatitis in adults. J Invest Dermatol 139:2090-7.

Tabachnick BG, Fidell LS (2013) Using multivariate statistics (6th ed.). New Jersey, Pearson Education, s. 60-113.

Tatar A, Saltukoglu G (2010) CES-Depresyon Ölçeği'nin Doğrulayıcı Faktör Analizi ve Madde Cevap Kuramı Kullanımı ile Türkçe'ye Uyarlanması ve Psikometrik Özelliklerinin İncelenmesi. Klinik Psikofarmakol Bülteni 20:213-27.

The Euro Qol Group (1990) EuroQol – a new facility for the measurement of health related quality of life. Health Policy 16:199-208.

Wang X, Guo G, Zhou L et al (2017) Health-related quality of life in pregnant women living with HIV: a comparison of EQ-5D and SF-12. Health Qual Life Outcomes 15:158.

Ware JE, Snow KK, Kosinski M et al (1993) SF-36 Health Survey: Manual and Interpretation Guide. Boston: The Health Institute, New England Medical Center.

Ware JE, Kosinski M, Keller SD (1994) SF-36 Physical And Mental Health Summary Scales: A User's Manual. Boston: The Health Institute, New England Medical Center.

Ware JE, Kosinski M, Keller SD (1995) SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston: The Health Institute, New England Medical Center.

Ware Jr JE, Kosinski M, Keller SD (1996) A 12-item Short-Form Health Survey. Construction of scales and preliminary tests of reliability and validity. Med Care 34:220–33.

Webster KE, Feller JA (2016) Comparison of the short form-12 (SF-12) health status questionnaire with the SF-36 in patients with knee osteoarthritis who have replacement surgery. Knee Surg Sports Traumatol Arthrosc 24:2620-6.

Wells GA, Russell AS, Haraoui B et al (2011) Validity of quality of life measurement tools-from generic to disease-specific. J Rheumatol (Suppl. 88):2-6.

Wójcik G, Zawisza K, Jabłońska K et al (2019) Transition out of Marriage and its Effects on Health and Health–Related Quality of Life among Females and Males. Courage and Courage-Polfus–Population Based Follow-Up Study in Poland. Appl Res Qual Life 1: 1-37.

Wukich DK, Sambenedetto TL, Mota NM et al (2016) Correlation of SF-36 and SF-12 component scores in patients with diabetic foot disease. J Foot Ankle Surg 55: 693-6.

Younsi M (2015) Health-related quality-of-life Measures: evidence from Tunisian population using the SF-12 health survey. Value Health Reg Issues 7: 54-66.

Younsi M, Chakroun M (2014) Measuring health-related quality of life: psychometric evaluation of the Tunisian version of the SF-12 health survey. Qual Life Res 23: 2047-54.

Yu DSF, Yan ECW, Chow CK (2015) Interpreting SF-12 mental component score: an investigation of its convergent validity with CESD-10. Qual Life Res 24: 2209-17.