J Surg Med. 2018;2(1):11-16.
DOI: 10.28982/josam.368341

Research article
Araştırma makalesi

Journal of Surgery and Medicine

e-ISSN: 2602-2079

Reliability and validity of the Turkish version of short form 36 (SF-36) in patients with rheumatoid arthritis

Romatoid artritli hastalarda Türkçe kısa form 36'nın (SF-36) güvenilirlik ve geçerliliği

Başak Bilir Kaya [1], Afitap İçağasıoğlu [2]

[1] Erenköy Physical Therapy and Rehabilitation Hospital, Istanbul, Turkey
[2] Medeniyet University, Goztepe Training and Research Hospital, Department of Physical Therapy and Rehabilitation, Istanbul, Turkey

Corresponding author / Sorumlu yazar:
Başak Bilir Kaya
Address / Adres: 19 Mayıs Mah. Şemsettin Günaltay Cad. Sultan Sok No:14, Erenköy Kadıköy, İstanbul, Türkiye
E-mail: basakbilir@gmail.com

Ethics Committee Approval is taken from Ethical Committee of Medeniyet University, Goztepe Training and Research Hospital at 20.10.2006, with record number of 32/P
Etik Kurul Onayı Medeniyet Üniversitesi, Göztepe Eğitim ve Araştırma Hastanesi Etik Komite'sinden 20.10.2006 tarih ve 32/P karar no ile alınmıştır

Conflict of Interest: No conflict of interest was declared by the authors.
Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemişlerdir.

Financial Disclosure: The authors declared that this study has received no financial support.
Finansal Destek: Yazarlar bu çalışma için finansal destek almadıklarını beyan etmişlerdir.

Received / Geliş Tarihi: 18.12.2017
Accepted / Kabul Tarihi: 05.02.2018
Published / Yayın Tarihi: 05.02.2018

Abstract

Aim: Investigating reliability and validity of the Turkish version of short form-36 (SF-36) in patients with rheumatoid arthritis

Methods: Demographic data of the patients with rheumatoid arthritis were recorded. Health Assessment Questionnaire (HAQ) and Short Form 36 (SF-36) were filled out. Disease activities were computed using Disease Activity Score 28 (DAS-28). Patients were recalled after three months and were asked to state how they felt compared to their first visit, and the same tests were repeated.

Results: 141 patients were admitted (9.9% male, 90.1% female). In the reliability study of SF-36, the Cronbach alpha value of the subscales varied in the range 0.792-0.992, hence SF-36 was found to be highly reliable. The item total score correlations were computed for each subscale and were found to be in the ranges: 0.436-0.840 for physical functioning, 0.887-0.895 for role function (physical), 0.861-0.958 for pain, 0.564-0.892 for general health perception, 0.702-0.841 for vitality (energy/fatigue), 0.949-0.952 for social functioning, 0.396-0.473 for role function (emotional) and 0.456-0.824 for mental health. The SF-36 scores from two consecutive visits spaced 3 months apart were compared and the p values were found to be greater than 0.05. The validity study was conducted for the 63 patients whose reported conditions did not change between two visits. The test-retest relation was evaluated using intra-class correlation coefficients, which ranged from 0.51 to 0.78 and the correlations of the two tests were found to be statistically significant. The comparison of SF-36 scores from two consecutive visits, all with $p>0.05$, showed no statistically significant changes.

Conclusion: The Turkish version of SF-36 was found to be reliable and valid in patients with rheumatoid arthritis.

Keywords: Reliability, Validity, Rheumatoid arthritis, SF-36

Öz

Amaç: Romatoid artritte Türkçe SF-36'nın güvenilirlik ve geçerliliğini araştırmak.

Yöntemler: Çalışmaya alınan Romatoid artritli hastaların demografik özellikleri kayıt edildi. Sağlık değerlendirme anketi (HAQ) ve SF-36 formları doldurduldu. Hastalık aktiviteleri DAS-28 kullanılarak hesaplandı. Hastalar 3 ay sonra tekrar çağırılarak kendilerini ilk gelişlerine göre nasıl hissettikleri soruldu ve ilk gelişlerinde yapılan testler tekrarlandı.

Bulgular: Çalışmadaki 141 hastanın %9,9'u erkek; %90,1'i kadındır. SF-36'nın güvenilirlik çalışmasında ölçek alt boyutlarının Cronbach alfa değerleri 0,792-0,992 arasında değişmekte olup, bunun sonucunda SF-36 yüksek düzeyde güvenilir bulunmuştur. Madde-toplam puan korelasyon katsayıları da her bir alt ölçek için ayrı ayrı hesaplanmıştır. Fiziksel fonksiyonda 0,436-0,840, fiziksel rol güçlüğünde 0,887-0,895; ağrıda 0,861-0,958, sağlığın genel olarak algılanmasında 0,564-0,892; vitalite (enerji)'de 0,702-0,841; sosyal fonksiyonda 0,949-0,952; emosyonel rol kısıtlamasında 0,396-0,473 ve mental sağlıkta 0,456-0,824 arasında bulunmuştur. Daha sonra her iki SF-36 ölçümü karşılaştırılmış ve $p>0,05$ olduğu için ikinci ölçümlerde ilk ölçümlere göre bir farklılık saptanmamıştır.

SF 36 geçerlilik çalışması, beyana dayalı durumları ikinci gelişlerinde ilkine göre değişmeyen 63 olgu üzerinde yapılmış olup; test tekrar test arasındaki korelasyon, intra-class korelasyon katsayıları ile değerlendirildiğinde her iki uygulama arasında istatistiksel olarak anlamlı ilişki görülmektedir. Bu çalışmada korelasyon katsayıları 0,51 ile 0,78 arasında değişmektedir. SF-36 puanlarının ilk ve ikinci ölçümleri arasında istatistiksel olarak anlamlı farklılık görülmemektedir ($p>0,05$).

Sonuç: Sonuç olarak Türkçe SF-36 Romatoid artritli hastalarda güvenilir ve geçerli bulunmuştur.

Anahtar kelimeler: Güvenilirlik, Geçerlilik, Romatoid artrit, SF-36

How to cite / Atıf için: Kaya BB, İçağasıoğlu A. Reliability and validity of the Turkish version of short form 36 (SF-36) in patients with rheumatoid arthritis. J Surg Med. 2018;2(1):11-16.

Introduction

Patients with rheumatoid arthritis often report symptoms that impair their quality of life. There is no clear idea of what is the best measure to use for the assessment of this chronic disease, and the factors that most of the scales take into account are limited to acute phase reactants such as joint involvement and sedimentation (ESR). Many researchers also include the disability part of the health assessment questionnaire (HAQ) to be more comprehensive [1]. HAQ currently also used to assess physical disability in rheumatoid arthritis. Its' validation study was carried out in many countries including England and Turkey [2]. It is often used in the United States to assess the outcome of observational and clinical trials [3,4]. One study showed that the most commonly used health status questionnaire in patients with outpatient rheumatology clinics is HAQ, although it is not known exactly how much it affects treatment decisions [5]. There is increasing evidence that HAQ is predictive for progression of the disease in the long term [6,7] and short term [8]. Although it is a valuable measure of physical function losses caused by RA, it does not include emotional and psychological problems. It cannot show exactly how the patient perceives his illness causes loss of quality of life. In addition, HAQ attaches more importance to the functions of the upper limb than to evaluate all areas of the physical function. It is particularly important to use health assessment questionnaire that measures quality of life in chronic illnesses. Ideally, this health assessment questionnaire should include general health information that is valid (accurate measurement of expected outcome) and reliable (which may give similar results when repeated), acceptable to patients and clinicians, applicable in intensive clinical conditions, and specific for the disease [9, 10]. SF-36 (short form 36) [11] is a general health screening survey designed in the United States. It has been shown to be valid and acceptable in a healthy population and reliability studies have been performed in different patient groups [12]. Koçyiğit et al. [13] translated SF 36 into Turkish and evaluated the reliability and validity of 50 patients with osteoarthritis and 50 patients with low back pain. Similar to HAQ, the patient-centered approach takes the patient's ill effects into the patient rather than the disease and blood test-focused outcomes used by the clinician. It differs from HAQ because the questions are designed to examine eight different subgroups of health and include physical activity constraints that the patient perceives as disease-related. However, there is no study on the reliability and validity of Turkish SF-36 in RA patients. The aim of this study is to assess the reliability and validity of Turkish SF-36 in RA patient group.

Materials and methods

One-hundred-forty-one rheumatoid arthritis patients who admitted to outpatient clinic of department of physical therapy and rehabilitation of Goztepe training and research hospital were taken into the study.

Inclusion criteria were designated as: patient older than 18 years old and younger than 90 years old, having RA diagnosis according to the 1987 ACR criteria. Exclusion criteria were designated as: patients with advanced heart failure (grade 3-4), acute or chronic kidney failure, and malignancy.

The health assessment questionnaire (HAQ) and SF-36 were used to measure quality of life after the demographic characteristics of the patients were recorded. The Health Assessment Questionnaire (HAQ) is composed of 20 questions and questions eight activities. They include dressing and grooming, arising, eating, walking, hygiene, reach and grip. Each answer is rated from 0 to 3 [14].

The DAS28 score in the evaluation of disease activity was calculated as follows.

DAS28 = 0.56 x sqrt(tender28) + 0.28 x sqrt(swollen28) + 0.70 x ln (ESR) + 0.014 x GH

There are special types of calculators and software that make these calculations and such software is used in calculations. The value obtained was used to classify the disease activity as follows; the patient is in remission at 2. 4 or below; between 2.4 and 3.6: low disease activity; between 3.6 and 5.5: moderate disease activity; over 5.5: high disease activity [15].

The hourly sedimentation value used to assess disease activity was noted from the routine laboratory tests performed by the patients and the Rheumatoid Factor was taken as positive or negative evaluation from any test performed during or after the diagnostic phase. Patients were recalled after 3 months and asked how they felt themselves based on their first arrival. Then they are divided to three groups according to their answers: better, same and worse. Then the tests that were carried out at the first arrival were repeated.

Statistical Analysis

For the statistical analyzes, the NCSS 2007 & PASS 2008 Statistical Software (Utah, USA) program was used. When the study data were evaluated, descriptive statistical methods (mean, standard deviation (SD), median) as well as Paired sample t test were used for intra-group comparison of normal distribution parameters and Wilcoxon sign test was used for intra-group comparison of non-normal distribution parameters. Mc-Nemar test and Cohen Kappa coherence analysis were used for the comparison of qualitative data. Spearman's correlation analysis and Intra-class Correlation Coefficient were used to evaluate the inter-parameter relations. The results were assessed at a 95% confidence interval (CI) and a significance level of $p < 0.05$.

Results

Demographic Characteristics

The study was carried out on a total of 141 cases between the dates of 01.09.2006 - 01.04.07. Patents' ages ranged from 22 to 80 years. 9.9% of the patients (n = 14) were male; 90.1% (n = 127) were females. The additional diseases and habits of the cases are shown in Table 1 and the distribution of drugs used is shown in Table 2.

Disease Activity Measurements

When we evaluated the activation classification according to the DAS 28 total score by McNemar test at the 2nd measures according to the first measures, we found that there was no significant difference between them with the significance close to each other ($p>0.05$).

Table 1: The Additional diseases and habits

	n	%
Additional Disease	59	41.8
Hepatitis B	2	1.4
Tuberculosis	10	7.1
Hypertension	33	23.4
Coronary Disease	14	9.9
Diabetes	16	11.3
Alcohol Abuse	2	1.4
Tobacco Use	18	12.8

Table 2: Drugs used by the patients

	n	%
Methylprednisolone	59	41.8
Prednisolone	2	1.4
Methotrexate	10	7.1
Chloroquine	33	23.4
Hydroxychloroquine	14	9.9
Lenflunamide	16	11.3
Sulfasalazine	2	1.4
Anti TNF	18	12.8
Nonsteroid anti-inflammatory	134	95.0
Bisphosphonate	134	95.0

Table 3: The values of the DAS 28 first and second measurements

	1st measurement Mean±SD	2nd measurement Mean±SD	p[1]
DAS 28.1	3.23±2.19 (3)	3.44±2.16 (3)	0.642
DAS 28.2	1.42±2.21 (0)	1.05±1.85(0)	0.016*
DAS 28.3	2.38±2.77 (2)	1.67±2.45 (1)	0.001**
DAS 28.4	22.10±15.49(18)	26.58±18.97 (22)	0.002**
DAS 28.5	3.49±11.95 (0.69)	1.69±5.45 (0.54)	0.140
DAS 28.6	3.37±2.17 (3)	3.35±2.12 (3)	0.397
DAS 28.7	1.94±0.91 (2)	1.53±0.92 (1)	0.001**
Total score	4.09±2.35 (3.39)	3.64±2.18 (2.92)	0.018*

[1] Wilcoxon Signed Rank test, *p<0.05, **p<0.01

DAS28.1: VAS (visual analog scale pain)
DAS28.2: Swollen joint (SW)
DAS28.3: Tender joint (TEN)
DAS28.4: Sedimentation (ESR)
DAS28.5: CRP
DAS28.6: The patients' general health (GH)
DAS28.7: Global disease activity (By doctor)

The non-coincidence rate is 32.6% (Cohen Kappa: 0.364). In the first measurement, the remission rate was 25.6% while in the second measurement it was 30.2%. low activity increased by 27.1% in the first measurement and 32.6% in the second measurement; at the first measurement, the moderate activity was seen as 22.5% in the second measurement and as 21.7% in the first measurement, the high activity rate was 24.8% in the first measurement and it decreased to 15.5% in the second measurement.

The values of the DAS 28 first and second measurements based on sub-parameters are presented in Table 3. When we evaluate the relationship between the DAS 28 initial measurements and the second measurements according to Spearman's correlation analysis, there is a significant correlation between the two measurements (p<0.01).

SF-36 Scale Validity and Reliability Analysis

The Cronbach alpha coefficient was used for the Validity and Reliability Analysis. Cronbach alpha coefficient is a measure of the similarity and closeness of individual points to each other in cases where individual points are found by collecting answers to questions on a scale containing k questions. The alpha coefficient is used to question whether the problem on the scale has formed a whole to explain or question a heterogeneous structure [16].

The evaluation of the alpha coefficient is based on the following criteria:

- If 0.0 <0.40, the scale is not reliable.
- If 0.40 <0.60 the scale has low reliability
- If 0.60 <0.80 is quite reliable
- If 0.80 <1.00, the scale is a highly reliable one

The SF-36 scale consists of 8 sub-dimensions. The subscales of the SF-36 questionnaire were validated for reliability. The Cronbach's alpha coefficients of the SF-36 subscales are shown in Table 4.

Table 4: The Cronbach's alpha coefficients of the SF-36 subscales

	Number of Question	Cronbach alpha coefficient
Physical Functioning	10	0.889
Physical Role Functioning	4	0.992
Pain	2	0.792
General Health	5	0.818
Vitality(energy)	4	0.816
Social Function	2	0.908
Emotional Role Functioning	3	0.836
Mental Health	5	0.918

Table 5: SF-36 measurements evaluation (n=129)

	1st measurement Mean±SD	2nd measurement Mean±SD	p[1]
Physical Functioning	68.83±20.97	72.17±18.38	0.065
Physical Role Functioning	55.03±44.56	58.52±45.33	0.391
Pain	58.64±22.10	59.79±22.24	0.555
General Health	49.44±21.56	47.32±19.78	0.215
Vitality(energy)	43.79±19.23	43.52±17.64	0.862
Social Function	72.28±26.23	76.35±23.55	0.107
Emotional Role Functioning	54.52±46.01	56.84±47.39	0.616
Mental Health	55.47±18.46	57.98±18.02	0.112

[1] Paired Samples t-test

Cronbach alpha values of scale subscales ranged from 0.792 to 0.992 and our scale was found to be highly reliable. The item-total score correlation coefficients for each subscale were calculated separately for the related items. It was found between 0.436-0.840 in physical function, 0.887-0.895 in physical role difficulty, 0.861-0.958 in pain, 0.564-0.892 in general perception of Health, 0.702-0.841 in Vitality (energy), 0.949-0.952 in social function, 0.396-0.473 in emotional role restriction, and 0.456-0.824 in mental health.

The evaluations made by collecting sub-dimension scores for both measurements are shown in Table 5. In the first measurement, 141 cases were applied; in the second measure this number dropped to 129. Our evaluations were made on 129 cases. There was no statistically significant difference between the changes of physical function, physical role function, pain, general health, vitality, social function, emotional role function and mental health scores after the first measurement (p>0.05).

In their second visit, the patients were asked about their previous development and the following results were obtained: the patients were worse by 17.1%, while 48.82% remained the same, and 34.1% were better. The SF-36 validity study will be based on 63 cases that have not changed and remained the same.

As shown in Table 6, there was no statistically significant difference between the physical function, physical

role function, pain, general health, Vitality, social function, emotional role and changes in mental health scores after the first measurement (p>0.05). The rise in social function score was found to be at the limit of significance.

Table 6: Evaluation of SF36 (n=63)

	1st measurement Mean±SD	2nd measurement Mean±SD	p[1]
Physical Functioning	69.36±19.80	72.69±17.36	0.154
Physical Role Functioning	58.73±43.10	64.28±44.61	0.362
Pain	61.30±21.49	63.79±18.26	0.313
General Health	49.50±21.29	49.52±16.79	0.995
Vitality(energy)	41.98±19.47	43.65±18.07	0.406
Social Function	74.20±25.18	80.75±20.05	0.050
Emotional Role Functioning	62.96±44.03	60.31±46.70	0.685
Mental Health	54.03±18.66	57.64±17.22	0.116

[1] Paired Samples t-test

SF 36 scale scores were evaluated on 63 patients who did not change at first visit according to their first visit. Correlation between test and re-test were evaluated with intra-class correlation coefficients and statistically significant relationship is observed between the two applications (Table 7). The correlation coefficients ranged from 0.51 to 0.78. The highest difference in the subsequent measures according to the first measure is seen in the social function score with 6.54 points.

Table 7: SF 36 Test re-test validity results (n=63)

SF 36 (0-100)	Intra-Class Correlation Coefficients	SF 36 scores Median Difference	SD	%95 CI
Physical Functioning	0.681	-3.33	18.3	-7.94-1.27
Physical Role Functioning	0.572	-5.55	48.03	-17.65-6.54
Pain	0.689	-2.49	19.43	-7.38-2.40
General Health	0.636	-0.01	19.82	-5.00-4.97
Vitality(energy)	0.785	-1.66	15.81	-5.64-2.31
Social Function	0.518	-6.54	25.97	-13.08-0.001
Emotional Role Functioning	0.523	2.64	51.60	-13.35-15.64
Mental Health	0.664	-3.61	18.01	-8.15-0.91

The comparison of two measurements of HAQ scores was made with Wilcoxon Signed Rank test and no statistically significant difference was observed (p>0.05). There was a statistically significant difference between the first physical function scores and HAQ 1 scores on the negative side and statistically 85.3% (p<0.01). Physical role score with a score of 60.1%, pain score was 61.7%; with a general health score of 42.5%, vitality score of 40.5%, social function score of 55.2%, emotional role function score of 34.9% and a mental health score of 29.6% in the negative direction (Table 8).

There was a statistically significant correlation (81.0%) between the second physical function scores and HAQ 2 scores on the negative side (p <0.01). Physical role score with a score of 62.5% with pain score at 65.9%; with a general health score of 45%, with vitality score of 48.7%, social function score of 51.3%, emotional role function score of 46.6% and a mental health score of 44.9% in the negative direction (Table 8).

Table 8: HAQ and SF 36 scores relation

	HAQ 1st measurement (n=141)		HAQ 2nd measurement (n=129)	
SF 36 (0-100)	r	p	r	p
Physical Functioning	-0.853	0.001**	-0.819	0.001**
Physical Role Functioning	-0.601	0.001**	-0.625	0.001**
Pain	-0.617	0.001**	-0.659	0.001**
General Health	-0.425	0.001**	-0.450	0.001**
Vitality(energy)	-0.405	0.001**	-0.487	0.001**
Social Function	-0.552	0.001**	-0.513	0.001**
Emotional Role Functioning	-0.349	0.001**	-0.466	0.001**
Mental Health	-0.296	0.001**	-0.449	0.001**

r: Spearman's correlation coefficient,** p<0.01

There was a statistically significant correlation (57.4%) between the first physical function scores and DAS 28 first measurement total scores on the negative side (p <0.01). The physical role was 53.8%; pain score at 57.5%; with a general health score of 23.4%, with vitality score of 17.9%, social function score of 46.6%, emotional role function score is 29.6%, mental health score is 17.0% all in negative direction (Table 9).

Table 9: DAS 28 total score and SF 36 scores relation

	DAS 28 1st measurement (n=141)		DAS 28 2nd measurement (n=129)	
SF 36	r	p	r	p
Physical Functioning	-0.574	0.001**	-0.543	0.005**
Physical Role Functioning	-0.538	0.001**	-0.571	0.001**
Pain	-0.575	0.001**	-0.653	0.001**
General Health	-0.234	0.005**	-0.365	0.001**
Vitality(energy)	-0.179	0.034*	-0.415	0.001**
Social Function	-0.466	0.001**	-0.413	0.001**
Emotional Role Functioning	-0.296	0.006**	-0.359	0.00**
Mental Health	-0.170	0.035*	-0.315	0.001**

r: Spearman's correlation coefficient,* p<0.05, ** p<0.01

A statistically significant correlation was found between the second physical function scores and DAS 28 total scores on the negative side 54.3% (p <0,01). Physical role score was 57.1%; pain score 65.3%; a general health score f 36.5%, 41.3% in vitality score, 41.3% in social function score; emotional role function score was 35.9% and 31.5% in the mental health score and all with negative direction (Table 9).

Discussion

The SF-36, commonly used in health care systems in the United States and UK, is a generic scale for measuring the health status and quality of life. Jenkinson et al. [17] stated that the performance of SF-36, especially psychometric and clinical validity, could be affected by the patient group and clinical condition under which the test was performed, and stated that satisfactory performance was not guaranteed for all patient groups and clinical conditions. This prediction requires that reliability and validity studies be conducted in that patient group before the scale is used in specific studies for specific patient groups. There is evidence that SF-36 is as good as disease-specific health scales in studies performed on some patient groups, for example in hip replacement patients [18,19]. In some subsequent studies, the validity of English SF-36 in the RA patient group [19,20] and responsiveness [21] was examined., Koçyiğit et al. [13] translated SF 36 into Turkish and conducted a reliability and validity study in 50 patients with osteoarthritis and 50 patients with low back pain.

Reliability shows the reproducibility of the accuracy of a measurement made by a scale [22]. The smaller the measurement error of your scale, the more reliable it is [22]. Internal consistency shows the relationship between a scale and the substance, and the extent to which the materials measure the desired concept. The classical statistical method is determined by the Cronbach-alpha coefficient (α). The α value is a number between 0 and 1, and the closer to one (1) is, the higher the internal consistency of the scale. Higher internal consistency supports the reliability of the scale [16].

This study was conducted with 141 patients with rheumatoid arthritis. Cronbach alpha coefficients of each subscale ranged from 0.772 to 0.992 in scale reliability studies, and according to these results, SF-36 is a reliable measure in patients with rheumatoid arthritis.

Developing the scale Ware et al. [23] Cronbach found the alpha coefficient between 0.62 and 0.94. Brazier et al. [24] they found Cronbach Alpha coefficients of sub-scales between 0.73 and 0.96. In a study of healthy individuals working in the UK, these values ranged from 0.76 to 0.90 [25]. In a study of patients with psoriatic arthritis, these values were obtained between 0.82-0.92 [26]. In a study of patients with systemic lupus erythematosus, Cronbach alpha values were calculated over 0.71 [27]. The Cronbach alpha coefficients obtained in our study were found to be sufficient in terms of reliability. The item-total score correlation coefficients are generally between 0.396 and 0.958, and the relevant items for each subscale are calculated separately. Between 0.436 and 0.840 in physical function, between 0.887 and 0.895 in physical role strength; between 0.861 and 0.958 for pain, between 0.564 and 0.892 for general perception of health; in vitality (energy) between 0.702 and 0.841; social functioning was between 0.949 and 0.952, emotional role restriction was between 0.396 and 0.473 and mental health was found between 0.456 and 0.824. In the study in which the scale was developed, item-total score correlation coefficients were found between 0.43-0.90 [23]. Husted et al. [26] found that the total score correlation coefficients were over 0.4 in their study of patients with psoriatic arthritis. Another study was performed in patients with systemic lupus erythematosus and the coefficients were again found to be over 0.4 [27].

All the correlation coefficients obtained in our study are significant and are higher than the correlations between the sub-scales in which the substances do not belong. As a result, SF-36 was found to be reliable for Turkish patients with rheumatoid arthritis.

Another process for the reliability of SF-36 is the test-again test evaluation. There was no statistically significant difference between the physical function, physical role difficulties, pain, general health, Vitality, social function, emotional role, and mental health scores after the initial measurement ($p>0.05$) (Table 6).The difference between the physical function, physical role difficulties, pain, general health, Vitality, social function, emotional role and mental health scores after the initial measurement ($p>0.05$) (Table 6). In other words, SF-36 has the test-retest reliability in patients with rheumatoid arthritis.

The validity study was performed on 63 cases whose SF 36 scale scores were not changed according to their initial arrival. The relationship between the test and re-test is evaluated with the in-class correlation coefficient; there is a statistically significant correlation between the two applications. In our study, correlation coefficients ranged between 0.51 and 0.78 (Table 7). The highest difference in the subsequent measures according to the first measure was seen in the social function score with 6.54 points.

Birrell et al. [28] in a study of 86 rheumatoid arthritis patients in the UK, physical and social functions of SF-36 were correlated well with HAQ, while the physical and emotional role subgroups were independent of disease activity. Based on these results, it was emphasized that the physical and emotional role subgroups of SF-36 were not related to the main scales of disease activity.

In this study, when the first SF-36 measurements of patients were compared with the HAQ, it was found that there was a statistically significant in negative way (85.3%), when compared with the SF-36 subunits, the emotional role function score was 34.9% and in mental health subgroups % 29,6 in negative way (Table 8). This shows that the emotional role subgroup is not directly affected by disease activity in rheumatoid arthritis. There may be different aspects which effects emotional situation in rheumatoid arthritis patients other than disease activity. Furthermore, in a study of 60 patients with rheumatoid arthritis using infliximab, emotional role function and mental health subgroups were also shown to be the least susceptible to change [29].

In conclusion, Turkish SF-36 in rheumatoid arthritis group was found to be reliable and valid. As the quality of life in the definition of 'health' of the World Health Organization is also important, it is recommended to use SF-36 in patients with rheumatoid arthritis in order to determine the quality of life of the patients and to reveal the psychosocial aspect of the disease.

References

1. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137-45.
2. Kucukdeveci AA, Sahin H, Ataman S, et al. Issues in crosscultural validity: example from the adaptation, reliability and validity testing of a Turkish version of the Standford Health Assessment Questionnaire. Arthritis & Rheumatism; Arthritis Care & Research. 2004;51(1):14-9.
3. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol. 1986;25(2):206-209.
4. Leigh JP, Fries JF, Parikh N. Severity of disability and duration of disease in rheumatoid arthritis. J Rheumatol. 1992;19(12):1906-11.
5. Buchbinder R, Bombardier C, Yeung M, Tugwell P. Which outcome measure should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to treatment in a randomised controlled trial. Arthritis Rheum. 1995;38(11):1568-80.
6. Carr A, Thompson P, Young A. Do health status measures (HSM) have a role in rheumatology? A survey of the use of and attitudes towards health status measures in the UK. Arthritis Rheum. 1996; ACR Abstracts S261.
7. Wolfe F, Cathey MA. The assessment and prediction of functional disability in rheumatoid arthritis. J Rheumatol. 1991;18(11):1774.

8. Leigh JP, Fries JF. Predictors of disability in a longitudinal sample of patients with rheumatoid arthritis. Ann Rheum Dis. 1992;51(5):581-7.
9. Young A. Short-term outcomes in recent-onset rheumatoid arthritis. Br J Rheumatol. 1995;34(suppl. 2):79-86.
10. Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A, Mowat A. Importance of sensitivity to change as a criterion for selecting health status measures. Quality Health Care. 1992;1(2):89-93.
11. Liang MH, Jette AM. Measuring functional ability in chronic rheumatoid arthritis. A critical review. Arthritis Rheum. 1981;24(1):80-6.
12. Ware JE, Sherbourne CD. The MOS 36-item short-form health status survey (SF-36). 1. Conceptual framework and item selection. Med Care. 1992;30(6):473-83.
13. Koçyiğit H, Aydemir Ö, Ölmez N, et al. SF-36'nın Türkçe için güvenilirliği ve geçerliliği. İlaç ve Tedavi Dergisi. 1999;12:102-6.
14. Küçükdeveci A, Şahin H, Ataman Ş, Griffiths B, Tennant A. Issue in cross-cultural validity: example from the adaptation, reliability, and validity testing of a Turkish version of thr Stanford Health Assessment Questionnaire. Arthritis & Rheum. 2004;51(1):14-9.
15. Aletaha D, Ward MM, Machold KP, Nell VPK, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis. Defining criteria for disease activity states. Arthritis Rheum. 2005;52(9):2625-36.
16. Coronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334.
17. Jenkinson C, Peto V, Coulter A. Making sense of ambiguity: evaluation of internal reliability and face validity of SF-36 questionnaire in women presenting with menorrhagia. Qual Health Care. 1996;5(1):9-12.
18. Hurst NP, Kind P, Ruta DA, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol. 1997;36(5):551-9.
19. Kosinski M, Keller SD, Ware JE, Hatoum HT, Kong SX. The SF-36 health survey (SF-36) as a generic outcome measure in clinical trials of patients with osteoarthritis and rheumatoid arthritis: relative validity of scales in relation to clinical measures of arthritis severity. Med Care. 1999;37(5 suppl):23-39.
20. Talamo J, Frater A, Gallivan S, Young A. Use of the short form 36 (SF-36) for health status measurement in rheumatoid arthritis. Br J Rheumatol. 1997;36(4):463-9.
21. Kvein TK, Smestad LM, Uhlig T. The responsiveness of generic and disease specific health health status measures in 759 patients with rheumatoid arthritis (RA). Arthritis Rheum. 1996;39(suppl.):260.
22. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. New York: Oxford University Press; 1989.
23. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and interpretation guide. Boston: The Health Institute. New England Medical Center, 1993.
24. Brazier JE, Harper R, Jones NMB et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ. 1992;305(6846):160-4.
25. Jenkinson C, Coulter A, Wright L: Short Form-36 (SF-36) health survey questionnaire: Normative data for adults of working age. BMJ. 1993;306(6890):1437-40.
26. Husted JA, Gladman DA, Farewell VT et al. Validating the SF-36 health survey questionnaire in patients with psoriatic arthritis. J Rheumatol. 1997;24(3):511-7.
27. Stoll T, Gordon C, Seifert B et al. Consistency and validity of patient administered assessment of quality of life by the MOS SF-36; its association with disease activity and damage in patients with systemic lupus erythematosus. J Rheumatol. 1997;24(8):1608-14.
28. Birrell FN, Hassell AB, Jones PW, Dawes PT. How does the short form 36 health questionnaire (SF-36) in rheumatoid arthritis (RA) relate to RA outcome measures and SF-36 population values? A cross-sectional study. Clin Rheumatol. 2000;19(3):195-9.
29. Russell AS, Conner-Spady B, Mintz A, Maksymowych WP. The responsiveness of generic health status measures as assessed in patients with rheumatoid arthritis receiving infliximab. J Rheumatol. 2003;30(5):941-47.