Translation of the RAND 36-Item Health Survey 1.0 (aka SF-36) into Arabic

Saud Abdulaziz Al Abdulmohsin
Stephen Joel Coons
JoLaine R. Draugalis
Ron D. Hays

P-7995

Published 1997 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1333 H St., N.W., Washington, D.C. 20005-4792

Translation of the RAND 36-Item Health Survey 1.0 (aka SF-36) into Arabic

Saud Abdulaziz Al Abdulmohsin, MS[1], Stephen Joel Coons, PhD[2],

JoLaine R. Draugalis, PhD[2], Ron D. Hays, PhD[3]

[1] Pharmacy Services Division, Saudi ARAMCO Medical Services Organization, Saudi ARAMCO Company, Dhahran, Saudi Arabia

[2]Center for Pharmaceutical Economics and Department of Pharmacy Practice and Science, College of Pharmacy, The University of Arizona, Tucson, AZ

[3]Social Policy Department, RAND, Santa Monica, CA and Department of Medicine, College of Medicine, University of California, Los Angeles, CA

The authors wish to acknowledge the assistance of Saudi ARAMCO Company, The University of Arizona's Center for Pharmaceutical Economics, and RAND. At the time this study was conducted, Mr. Al Abdulmohsin was a graduate student in the Department of Pharmacy Practice at The University of Arizona College of Pharmacy.

Abstract

The objectives of this research were to: (1) translate the RAND 36-Item Health Survey 1.0 (aka SF-36) into Arabic; (2) evaluate the reliability and equivalence of the Arabic and English versions in a sample of Saudi Arabian citizens; and (3) assess the health status of a sample of Saudi Arabian citizens using both the Arabic and English versions. Forward and backward translation of the SF-36 with committee review was performed. Both the Arabic and English versions of the survey were administered to a convenience sample of bilingual (English and Arabic) Saudi citizens (N=415) at Saudi ARAMCO Company, Dhahran, Saudi Arabia. Internal consistency, equivalent-forms, and test-retest reliability were estimated for the Arabic and English versions of the survey. The results of the study provide support for the reliability and equivalence of both versions. Additional studies are needed with a representative sample of the general Saudi population to further assess the psychometric properties of the Arabic version.

In recent years, concern about cost containment in health care and interest in the impact of medical interventions on functioning and well-being have resulted in increased attention to the measurement of health-related quality of life (HRQOL).[1] Many treatment efforts are now directed beyond curing a disease or extending life to emphasizing the reduction of morbidity in societies where chronic diseases are prevalent.[2] As a result, the evaluation of medical interventions is no longer limited to morbidity, mortality, and the biomedical parameters that have traditionally been measured.[3] HRQOL assessment is being used in the process of direct patient care, clinical trials, program evaluation, and in monitoring the health status of populations.[4]

The World Health Organization (WHO) conceptualizes proper function as the core of health. The WHO's definition of health as a state of "complete physical, mental and social well-being, and not just the absence of disease" has become the foundation for many HRQOL measures.[5] Thus, HRQOL measures need to include multidimensional assessments of functional status and well-being.[6]

Most of the existing HRQOL measures were developed in the English language and intended for use in English-speaking countries.[7] There is clearly a need for measures for use in non English-speaking cultures, and demand has accelerated rapidly because of a growing number of large multinational trials.[8]

The RAND 36-Item Health Survey 1.0/SF-36

The RAND 36-Item Health Survey 1.0 is a generic health-related quality of life measure.[9] It is comprised of the same items included in the MOS 36-Item Short-Form Health Survey (SF-36).[10] Both measures emanated from work begun at RAND in 1984 as part of the Medical Outcomes Study (MOS). The only difference between the measures is the scoring algorithm applied when

calculating two of the scale scores (i.e., general health and pain.), which is independent of the measurement process. Hence, for the purpose of comparability with published general population norms, SF-36 scoring was used and the measurement instrument under investigation will be referred to as the SF-36.

The SF-36 includes multi-item scales that assess eight health concepts: physical functioning; bodily pain; role limitations due to physical health problems; role limitations due to emotional problems; emotional well-being (mental health); social functioning; energy/fatigue (vitality); and general health perceptions. A single item that provides an indication of perceived change in general health status over a one-year period is also included. The SF-36 has been translated into a number of languages, and psychometric testing of the translated versions provides evidence that the SF-36 is a reliable and valid general health survey measure across different cultures or nations.[11,12]

The original SF-36 was developed in the United States and is considered the U.S.-English version (since other English versions are now available for Canada, Australia, and the United Kingdom). However, in the remainder of this paper, the U.S.-English version of the SF-36 will be referred to as the English version for brevity. In this study, the SF-36 was translated into Arabic. The reliability and equivalence of the Arabic and U.S.-English versions of the SF-36 scales were assessed.

Methods

Translation Steps

1. Forward Translation

a. Six Saudi professionals independently translated the SF-36 into Arabic. Four of the

professionals were unaware of the background and purpose of the survey; however, the other two were briefed about the nature of the instrument and its purpose. This balance of naive and informed translators was designed to optimize the quality of the translation.

b. An additional translation of the SF-36 was performed by the Saudi ARAMCO Company Translation Unit in Dhahran, Saudi Arabia. Those involved in this translation were unaware of the background and purpose of the SF-36.

c. A committee composed of three of the six individuals involved in step 1a was established. Each committee member was informed about the background and purpose of the SF-36 health survey. The committee discussed all the translations produced in steps 1a and 1b and produced a new translation.

d. An additional translation was made by the first author (S. Al Abdulmohsin) after reviewing all translations and selecting the translation for each item and response scale with which the majority of translators agreed.

2. Back-Translation. A back-translation was performed independently for both forward translations obtained from steps 1c and 1d. In addition, the Arabic translation made by the Saudi ARAMCO Company Translation Unit (step 1b) was back translated into English by the same Translation Unit, but by individuals who were not involved in the forward translation. All those doing the back-translations were unaware of the background and purpose of the SF-36. Back-translators without prior knowledge of the original instrument's intent were used to minimize potential bias and their back-translations were expected to uncover unanticipated problems with language and interpretation.

3. Committee Review. A committee was formed comprised of seven bilingual members from

Saudi Arabia whose mother language is Arabic. The committee members represented different geographical regions in Saudi Arabia (north, west, east, south, and central) and had not been involved in previous translation steps. Regional diversity on the committee provided a comprehensive review of each Arabic word by individuals from different sub-cultures throughout the nation.

All committee members were Saudi graduate students working on their doctorates or master's degrees in different fields at The University of Arizona in Tucson. The committee reviewed all the translated material available. Their review included introductory material, item stems, and response choices for each item. Committee members also reviewed back-translations when it was necessary. Based on a consensus of the committee, a decision was made regarding the best translation. If a satisfactory translation was not available, the committee devised a substitute translation.

Four of the SF-36 items were modified from the U.S.-English version to be relevant to the Saudi culture. The physical functioning item that includes bowling and playing golf was changed to cleaning and taking care of one's garden, because bowling and golf are rare in Saudi Arabia. The physical functioning item that includes "a mile" was replaced by "kilometer and half" because the metric system is used in Saudi Arabia. "Several blocks" and "one block" were similarly replaced with "a half kilometer" and "a hundred meters," respectively, for two other physical functioning items because Saudi people do not use the word "block" to describe distance.

Pretest

A pretest was conducted to provide the opportunity to correct errors or potential problems before implementing the main study. A sample of Saudi students enrolled at The University of

Arizona and their spouses were selected to respond to the final translated Arabic version of the SF-36. Feedback was obtained from respondents regarding their overall under-standing of the questionnaire (i.e., the clarity of the instructions, items, and response scales). This feedback helped identify additional changes that needed to be made to improve the layout of the questionnaire, including closer placement of the question stem to the corresponding items. (The final version of the study instrument appears in the Appendix.)

Main study

Setting

The main study was conducted at Saudi ARAMCO Company in the eastern province of the kingdom of Saudi Arabia. Respondents were from ARAMCO locations in Dhahran, Al-Hasa, 'Udhailiyah, Abqaiq, and Ras-Tanura.

Sample

A convenience sample was selected from several departments at Saudi ARAMCO Company. This sample was selected by asking each of the departments to generate a list of bilingual (Arabic and English) Saudi employees who were willing to participate in the study. Persons were selected from each department to work as coordinators with the first author. The role of the coordinators was as follows: (1) distribute the questionnaires to the employee participants in their departments; (2) collect the questionnaires after they were completed by the participants; and (3) follow up with non-respondents to encourage completion.

After obtaining complete lists of participants from the departments, each person was assigned randomly to one of six groups (using a computer program developed for this purpose). Eighty-nine subjects were assigned to each of groups 1, 2, and 3, and 88 subjects were assigned to each of groups 4, 5, and 6.

Data Collection

Data were collected by distributing the English and/or Arabic versions of the SF-36 personally through the coordinators to the participants in January 1995 at Saudi ARAMCO Company in the eastern province of the kingdom of Saudi Arabia. Cover letters explaining the purpose and objective of the study were provided. Additional items were included to gather demographic information (i.e., gender, age, and education) and the questionnaires were coded so that non-respondents could be contacted. Only the first author had access to the list that linked the questionnaire codes to specific individuals. These questionnaires were administered to the six groups two times (time 1= initial; time 2= retest) as follows:

Group 1. The participants in Group 1 completed the English version immediately followed by the Arabic version. The assigned coordinator distributed a sealed envelope containing both versions to the participants. The participants were instructed to complete both questionnaires as follows: (a) complete the English version first, (b) insert the English questionnaire after completion in the envelope provided and seal it, (c) complete the Arabic questionnaire then insert it in the envelope and seal it, and (d) personally return both envelopes to the assigned coordinator.

Group 2. The participants in Group 2 completed the Arabic version first, immediately followed by the English version. Aside from the order of administration, instructions for Groups 2 through 6 were as discussed for Group 1.

Group 3. The participants in Group 3 completed the English version followed by the Arabic version after a two-week interval.

Group 4. The participants in Group 4 completed the Arabic version followed by the English version after a two-week interval.

Group 5. The participants in Group 5 completed the Arabic version twice (at the first time point and two weeks later).

Group 6. The participants in Group 6 completed the English version twice (at the first time point and two weeks later).

All participants were instructed to return the completed questionnaires to the assigned coordinators using the envelopes provided.

Scoring

All scales were transformed linearly to a 0 to 100 possible range of scores, with 0 representing the least favorable health state and 100 representing the most favorable health state. All scores reflect the percent of the total possible score for that scale. For comparison of scores from this sample to that of the U.S. general population, scoring of the SF-36 followed the procedures used for the published norms.[13] For other analyses, scoring of the pain and general health perception scales followed the RAND recommendations.[9] There is no evidence for the superiority of one scoring method over the other, but the RAND method is more straightforward and consistent across items and scales.

Data Analysis

Statistical analyses were conducted using SPSS for VAX/VMS, Version 4.1.

Descriptive Statistics. The mean ages, percent male and female, and education level were calculated. Mean SF-36 scale scores were also computed.

Internal Consistency Reliability. Cronbach's coefficient alpha [14] was used to estimate internal consistency reliability coefficients for the initial Arabic administrations (Groups 2, 4, and 5), Arabic retest administrations (Groups 1, 3, and 5), initial English administrations (Groups 1, 3,

and 6), and English retest administrations (Groups 2, 4, and 6).

Test-Retest Reliability. Pearson product-moment correlations were computed between initial and retest administrations in Groups 5 and 6 to assess test-retest reliability over a two-week interval.

Equivalent-Forms Reliability. Pearson product-moment correlation coefficients were used to estimate equivalent forms reliability between initial and retest administrations in Groups 1 and 2.

Scale Score Means. Scale score means were computed for the Arabic and English versions of the SF-36 to determine if central tendency varied by version. In addition, paired t-tests were computed comparing initial and retest administrations.

Results

Subjects

Of the 531 subjects randomly assigned to the six study groups, 446 (84%) individuals returned their questionnaires. Of these, 31 individuals were excluded from the study because they had missing data for one or more item(s). Thus, the analytic sample size was 415.

The mean age of the subjects was 34.86 (SD=7.0) with a range of 19 to 59 (Table 1). Fifty-two percent of the sample were college graduates. A one-way analysis of variance comparing the age of the six groups revealed no statistically significant differences [$F(5, 358) = 0.929$, $p > 0.05$]. Chi-square analysis revealed no significant differences in the proportion of males and females in the six study groups.

Internal Consistency Reliability

Arabic Version

Initial Administration. Table 2 provides alpha reliability coefficients for the initial

administration of the Arabic version of the SF-36 (Groups 2, 4, and 5, n=212). Alpha coefficients ranged from 0.60 (general health) to 0.87 (physical functioning). Except for the general health scale, reliability coefficients exceeded 0.70, satisfying Nunnally's standard of acceptable reliability for group-level studies.[15] Alpha coefficients from the Medical Outcomes Study (MOS) are provided for comparison in the last column of Table 2.

Retest Administration. Alpha coefficients for retest administrations (Groups 1, 3, and 5, n=181) ranged from 0.57 (general health) to 0.88 (physical functioning), and exceeded 0.70 for all scales except for general health.

English Version

Initial Administration. As shown in Table 3, alpha coefficients for the initial administrations of the English version of the SF-36 (Groups 1, 3 and 6, n=181) ranged from 0.56 (role limitations - physical, social functioning) to 0.89 (physical functioning). Alpha coefficients from the MOS are provided in the last column of Table 3.

Retest Administration. Alpha coefficients for the retest administrations (Groups 2, 4 and 6, n=179) ranged from 0.60 (social functioning) to 0.90 (physical functioning).

Test-retest reliability

Two-week test-retest product-moment correlations (see Table 4) for the SF-36 ranged from 0.29 (bodily pain) to 0.80 (emotional well-being) for the Arabic version (n=63) and from 0.46 (role limitations - physical) to 0.77 (general health perceptions) for the English version (n=52).

Equivalent-forms reliability

Equivalent-forms product-moment correlations (see Table 5) between corresponding scales for those administered the English version immediately followed by the Arabic version (n=53)

ranged from 0.78 (social functioning) to 0.91 (emotional well-being). Correlations for those administered the Arabic version immediately followed by the English version (n=65) ranged from 0.73 (energy/fatigue) to 0.92 (emotional well-being).

Mean scores for the Arabic and English versions of the SF-36

The mean scale scores for the initial administration of the Arabic version were computed for participants in Groups 2, 4, and 5 (n=170). The mean scale scores for the initial administration of the English version were computed for participants in Groups 1, 3, and 6 (n=177). For this analysis, the bodily pain and general health perception scales were calculated using the SF-36 scoring system used to create the general U.S. adult population norms (see Table 6). The mean scale scores for a sample of adults in the general U.S. population matched for age and gender to the Saudi Arabian samples are also provided.[13]

Differences between the mean scale scores

Paired t-tests were computed comparing mean scale scores for the Arabic and English versions of the SF-36. There were no significant mean differences for those administered the English version followed immediately by the Arabic version (n=53), supporting the equivalence of the two versions (Table 7). For those administered the Arabic version followed immediately by the English version (n=65), no significant differences were found for six of the eight scales; however, means for the energy/fatigue and social functioning scales were significantly different (Table 8). The average differences between the energy/fatigue and social functioning scale scores were only about 3 points on the 0 to 100 score range.

Discussion

The median internal consistency reliability coefficients for all administrations (Groups 1, 3, and 5) of the Arabic version of the SF-36 exceeded 0.70 for every scale except for the general

health perceptions scale (median alpha =0.59). The median internal consistency reliability coefficients for all administrations (Groups 2, 4, and 6) of the English version of the SF-36 exceeded 0.70. Therefore, the results of this study provide support for the reliability of the Arabic version of the SF-36 and are consistent with previous reliability estimates reported for the English version.[9,16,17]

Both the Arabic and English versions of the SF-36 tended to have internal consistency reliability coefficients equal to or above the acceptable standards for group comparisons. However, the Arabic version had higher median values than the English version, and this may have been due to cultural and language factors. Although all participants were bilingual, their mother language is Arabic. As a result, it is likely that the participants had a better understanding of the Arabic version than the English version, leading to more internally-consistent (reliable) responses. In addition, the Arabic version of the SF-36 was adapted to the Saudi culture whereas the English version was developed for the dominant U.S. culture.

Test-retest correlations over a two-week time interval were similar for the English and Arabic versions of the SF-36. The size of the correlations indicate a noteworthy degree of variation between initial and retest administrations. The timing of the study is one factor that may have contributed to the fluctuation in scale scores. The study was conducted in January when the season was changing from summer to winter in the eastern province of Saudi Arabia, and this change in the weather may have impacted upon participants' functioning and well-being. Many Saudi citizens suffer from colds, influenza, headache, bronchitis, and related symptoms in the winter. Due to possible real changes, test-retest reliability estimates need to be evaluated with caution, especially when studying a dynamic process such as health status or quality of life.[18]

The equivalence of the corresponding Arabic and English versions of the SF-36 scales was assessed using correlations and paired t-tests. Product-moment correlations ranged from 0.73 to 0.92 between corresponding scales. Only two out of 16 t-tests of paired means between the Arabic and English versions of the SF-36 were statistically significant. These results provide strong support for the equivalence of the Arabic and English versions of the SF-36.

When the mean scale scores for the Arabic version of the SF-36 were compared to those for the general U.S. population, a number of significant differences were found (Table 6). The Saudi citizens (N=170) had significantly lower mean scale scores for five of the eight scales when compared to the U.S. sample. In addition, one scale score mean (i.e., energy/fatigue) was significantly higher in the Arabic sample than in the U.S. sample.

When the mean scale scores for the English version of the SF-36 were compared to those for the general U.S. population, four of the eight scale score means were significantly different (Table 6). Again, only the mean score for the energy/fatigue scale was higher in the Arabic citizens (N=177) than in the U.S. sample. The U.S. sample had higher mean scores than the Arabic sample for the other scales in which significant differences were found (i.e., physical functioning, social function, and general health).

The results of this study provide support for the reliability of the Arabic version of the SF-36 and its equivalence with the English language version. However, this sample is not representative of the Saudi population. Thus, these results may not generalize to the Saudi population.

Additional studies need to be conducted to test the reliability and validity of the Arabic version using a well-designed sampling frame from the general Saudi population. Studies are also needed to evaluate the sensitivity of the Arabic version to variation in disease severity as well as

its responsiveness to the effects of medical treatments. Further research should also be conducted in other Arab cultures in order to examine to what extent this new Arabic version is applicable and to make any modifications to the instrument if necessary.

References

1. Coons SJ, Kaplan RM. Quality of life assessment: understanding its use as an outcome measure. *Hospital Formulary* 1993;28:486.

2. Kaplan RM, Coons SJ, Anderson JP. Quality of life and policy analysis in arthritis. *Arthritis Care Res* 1992;5:173.

3. Ware JE Jr. The assessment of health status. In: Aiken LH, Mechanic D, eds. *Application of Social Science to Clinical Medicine and Health Policy*. New Brunswick, NJ: Rutgers University Press, 1986:204.

4. Patrick DL, Erickson P. *Health Status and Health Policy: Allocating Resources to Health Care*. New York: Oxford University Press, 1993.

5. World Health Organization. *Constitution of the World Health Organization*. Geneva: WHO Basic Documents, 1948.

6. MacKeigan LD, Pathak DS. Overview of health-related quality-of-life measures. *Am J Hosp Pharm* 1992;49:2236.

7. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46:1417.

8. Berzon R, Hays RD, Shoemaker SA. International use, application and performance of health-related quality of life instruments. *Qual Life Res* 1993;2:367.

9. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ* 1993;2:217.

10. 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.

11. Aaronson NK, Acquadro C, Alonso J, et al. International quality of life assessment (IQOLA) project. *Qual Life Res* 1992;1:349.

12. Ren XS, Amick B, Zhou L, Gandek B, Ware JE. Chinese version of the SF-36 Health Survey: a review of translation and report on the psychometric testing results. *J Clin Epidemiol* (in press).

13. Ware JE Jr, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: Nimrod Press, 1993.

14. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297.

15. Nunnally JC. *Psychometric Theory*, Second edition. New York, NY: McGraw-Hill, 1978.

16. Kurtin PS, Davies AR, Meyer KB, DeGiacomo JM, Kantz ME. Patient-based health status measures in outpatient dialysis: early experiences in developing an outcomes assessment program. *Med Care* 1992;30:MS136.

17. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. psychometric and clinical tests of validity in measuring physical and mental constructs. *Med Care* 1993;31:247.

18. Coons SJ, Kaplan RM. Assessing health-related quality of life: application to drug therapy. *Clin Ther* 1992;14:850.

Table 1. Demographic characteristics of the overall analytic sample (N=415)

Patient Characteristics	Percent of Subjects (Number of Cases)
Gender	
Male	78 (324)
Female	21 (87)
Not Provided	1 (4)
Education	
Primary School	3.4 (14)
Intermediate	8.2 (34)
Secondary School	35.2 (146)
B.S.	42.7 (177)
M.S.	7.7 (32)
Ph.D.	1.9 (8)
Not Provided	1 (4)

Table 2. Cronbach's alpha coefficients for the Arabic version of the SF-36

Scale	Number of Items	Groups 2, 4, & 5 Arabic-Initial (N=212)	Groups 1, 3, & 5 Arabic-Retest (N=181)	MOS (N=2471)[1]
Physical Functioning	10	0.87	0.88	0.93
Role Limitations - Physical	4	0.84	0.80	0.84
Role Limitations - Emotional	3	0.75	0.70	0.83
Energy/Fatigue	4	0.78	0.82	0.86
Emotional Well-being	5	0.83	0.84	0.90
Social Functioning	2	0.75	0.79	0.85
Bodily Pain	2	0.81	0.80	0.78
General Health	5	0.60	0.57	0.78
Range		0.60-0.87	0.57-0.88	0.78-0.93

Group 1 = English then Arabic Immediately
Group 2 = Arabic then English Immediately
Group 3 = English then Arabic with a Two-Week Interval
Group 4 = Arabic then English with a Two-Week Interval
Group 5 = Arabic then Arabic with a Two-Week Interval
Group 6 = English then English with a Two-Week Interval

[1] Data is from baseline of the Medical Outcomes Study (Reference 9).

Table 3. Cronbach's alpha coefficients for the English version of the SF-36

Scale	Number of Items	Groups 1, 3, & 6 English-Initial (N=181)	Groups 2, 4, & 6 English-Retest (N=179)	MOS (N=2471)[1]
Physical Functioning	10	0.89	0.90	0.93
Role Limitations -Physical	4	0.56	0.76	0.84
Role Limitations - Emotional	3	0.66	0.74	0.83
Energy/Fatigue	4	0.58	0.61	0.86
Emotional Well-being	5	0.68	0.79	0.90
Social Function	2	0.56	0.60	0.85
Bodily Pain	2	0.70	0.73	0.78
General Health	5	0.61	0.68	0.78
Range		0.56-0.89	0.60-0.90	0.78-0.93

Group 1 = English then Arabic Immediately
Group 2 = Arabic then English Immediately
Group 3 = English then Arabic with a Two-Week Interval
Group 4 = Arabic then English with a Two-Week Interval
Group 5 = Arabic then Arabic with a Two-Week Interval
Group 6 = English then English with a Two-Week Interval

[1] Data is from baseline of the Medical Outcomes Study (Reference 9).

Table 4. Test-retest (2-week interval) product-moment correlations for SF-36 scales (Groups 5 and 6)

Scale	Group 5 (N=63)	Group 6 (N=52)
Physical Functioning	0.73	0.46
Role Limitations - Physical	0.48	0.46
Role Limitations - Emotional	0.55	0.58
Energy/Fatigue	0.69	0.63
Emotional Well-Being	0.80	0.74
Social Function	0.49	0.56
Bodily Pain	0.29	0.50
General Health	0.70	0.77
Range	0.29-0.80	0.46-0.77

Group 5 = Arabic then Arabic with a two-week interval
Group 6 = English then English with a two-week interval

Note: All correlations are statistically significant ($p<0.05$)

Table 5. Equivalent-forms reliability product-moment correlations for corresponding SF-36 scales (Groups 1 and 2)

Scale	Group 1 (N=53)	Group 2 (N=65)
Physical Functioning	0.81	0.90
Role Limitations - Physical	0.88	0.76
Role Limitations - Emotional	0.87	0.88
Energy/Fatigue	0.86	0.73
Emotional Well-Being	0.91	0.92
Social Function	0.78	0.80
Bodily Pain	0.85	0.77
General Health	0.90	0.90
Range	0.78-0.91	0.73-0.92

Group 1 = English then Arabic immediately
Group 2 = Arabic then English immediately

Note: All correlations are statistically significant ($p<0.01$)

Table 6. Mean SF-36 scale scores for the initial administrations of the Arabic and English versions of the SF-36 compared with the general U.S. population

		Mean (SD)				
Scale	No. of Items	Group 2, 4, and 5 initial Arabic administrations (N=170)	Norms for the general U.S. population [1]		Group 1, 3, and 6 initial English administrations (N=177)	Norms for the general U.S. population [1]
Physical Functioning	10	87.3 (17.1)**	91.3 (3.4)**		82.4 (20.6)****	91.7 (3.4)****
Role Limitations - Physical	4	84.3 (30.3)*	88.8 (3.6)*		90.3 (19.4)	89.2 (3.5)
Role Limitations - Emotional	3	76.3 (35.2)**	83.3 (2.2)**		82.3 (29.3)	83.3 (2.1)
Energy/Fatigue	4	67.6 (16.8)**	63.5 (2.7)**		68.1 (16.8)***	63.6 (2.6)***
Emotional Well-Being	5	75.2 (16.1)	75.0 (1.9)		75.6 (15.6)	75.0 (1.8)
Social Function	2	82.0 (19.3)**	86.1 (2.2)**		79.0 (20.7)****	86.0 (2.0)****
Bodily Pain	2	80.3 (21.9)	79.8 (3.33)		81.1 (19.0)	80.0 (3.4)
General Health	5	71.1 (16.1)****	76.9 (2.8)****		74.5 (15.1)*	77.2 (2.7)*

[1]Mean scale scores for the general U.S. population were matched for age and gender to the individual samples.

Significance of differences between the mean SF-36 scale scores for the sample and the general U.S. population:
* $p < .05$
** $p < .01$
*** $p < .001$
**** $p < .0001$

Table 7. Comparison (paired t-test) of mean (SD) scale scores for consecutive administrations of the English then Arabic version of the SF-36

Scale	Administration	Group 1 (N=53) Mean (SD)	p-value
Physical Functioning	Initial Retest	89.53 (11.86) 91.42 (10.11)	0.560
Role Limitations - Physical	Initial Retest	93.40 (17.10) 94.34 (17.43)	0.419
Role Limitations - Emotional	Initial Retest	81.76 (31.05) 81.76 (32.40)	1.000
Energy/Fatigue	Initial Retest	67.07 (16.88) 69.25 (17.85)	0.091
Emotional Well-Being	Initial Retest	76.60 (14.04) 76.45 (16.01)	0.871
Social Functioning	Initial Retest	80.66 (17.26) 79.72 (18.55)	0.569
Bodily Pain	Initial Retest	89.48 (13.88) 90.80 (13.10)	0.201
General Health	Initial Retest	72.83 (16.86) 71.79 (16.35)	0.310

Table 8. Comparison (paired t-test) of mean (SD) scale scores for consecutive administrations of the Arabic then English version of the SF-36

Scale	Administration	Group 2 (N=65)	
		Mean (SD)	p-value
Physical Functioning	Initial Retest	89.77 (13.90) 89.46 (13.32)	0.686
Role Limitations - Physical	Initial Retest	85.38 (28.60) 84.62 (27.13)	0.748
Role Limitations - Emotional	Initial Retest	73.33 (38.73) 73.85 (37.96)	0.829
Energy/Fatigue	Initial Retest	69.92 (16.19) 66.62 (17.44)	0.035*
Emotional Well-being	Initial Retest	75.14 (17.66) 76.00 (18.11)	0.326
Social Functioning	Initial Retest	81.73 (19.15) 78.46 (20.67)	0.040*
Bodily Pain	Initial Retest	85.58 (18.65) 83.27 (19.99)	0.164
General Health	Initial Retest	70.85 (18.00) 70.54 (18.98)	0.762

* = significant at $P < 0.05$

APPENDIX

Arabic Version of the RAND 36-Item Health Survey 1.0

استبيان صحي

الجنس☐ ذكر

☐ انثى

العمر_____سنة

المؤهل العلمي: ابتدائي ☐

اعدادي ☐

ثانوي ☐

بكالوريوس ☐

ماجستير ☐

دكتوراه ☐

من فضلك، أجب على كل الأسئلة الموجودة في هذا الاستبيان. في حالة عدم وضوح أي سؤال، أرجو اختيار أقرب اجابة لمفهومك للسؤال.

١- بصورة عامة، كيف ترى حالتك الصحية؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

☐ ممتازة

☐ جيدجدا

☐ جيدة

☐ لا بأس بها

☐ سيئة

٢- **مقارنة بعام مضى**، كيف تقيم حالتك الصحية **الآن** بصورة عامة؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

☐ أفضل بكثير مما كانت عليه قبل عام

☐ أفضل نوعا ما من العام الماضي

☐ تقريبا على ما هي عليه

☐ أسوأ نوعا ما من العام الماضي

☐ أسوأ بكثير مما كانت عليه قبل عام

٣- تتعلق البنود التالية بأنشطة يمكن ان تقوم بها خلال يومك العادي. **في الوقت الحالي**، الى اي مدى **تقيدك حالتك الصحية**:	(اختر اجابة واحدة وضع علامة ✓ تحت الاجابة المناسبة)		
	نعم تقيدني كثيرا	نعم تقيدني قليلا	لا تقيدني اطلاقا
أ) من ممارسة **الأنشطة الشاقة** مثل: الجري، حمل الاشياء الثقيلة او مزاولة الأنشطة الرياضية المجهدة جدا؟	☐	☐	☐
ب) من ممارسة **الأنشطة متوسطة الجهد**، كتحريك الطاولة او التنظيف باستخدام المكنسة الكهربائية او تنظيف حديقة المنزل والعناية بها ؟	☐	☐	☐
ج) من حمل المشتريات من البقالة او السوق المركزي (السوبرماركت)؟	☐	☐	☐
د) من صعود الدرج **لعدة ادوار**؟	☐	☐	☐
هـ) من صعود الدرج **لدور واحد فقط**؟	☐	☐	☐
و) من الانحناء او الركوع او السجود ؟	☐	☐	☐
ز) من المشي لأكثر من **كيلومتر ونصف**؟	☐	☐	☐
ح) من المشي لمسافة **نصف كيلومتر**؟	☐	☐	☐
ط) من المشي لمسافة **مئة متر**؟	☐	☐	☐
ي) من الاستحمام او ارتداء الملابس بنفسك؟	☐	☐	☐

الصحة الجسمية

٤- تتعلق البنود التالية (أ ، ب ، ج ، د) بالمشاكل التي يمكن ان تواجهـــــك خلال تأديتك لعملك او للأنشطة اليومية المعتادة نتيجة لحالتك الصحية الجسمية. **خلال الأسابيع الأربعة الماضية**، هل تسببت **حالتك الصحية الجسمية في:**	(اختر اجابة واحدة وضع علامة ✓ تحت الاجابة المناسبة) نعم	لا
أ) **التقليل من الوقت** الذي تقضيه في العمل او اي انشطة اخرى؟	☐	☐
ب) **التقليل** مما تود **انجازه** من العمل أو أي أنشطة أخرى؟	☐	☐
ج) **تقييدك في أداء نوع** معين من الأعمال أو أي أنشطة أخرى؟	☐	☐
د) **أن تجد صعوبة في** تأدية العمل أو أي أنشطة أخرى؟ (على سبيل المثال، احتجت الى جهد اضافي لتأديتها)	☐	☐

الصحة النفسية

٥- تتعلق البنود التالية (أ ، ب ، ج) بالمشاكل التي يمكن ان تواجهك خلال تأديتك لعملك او الانشطة اليومية المعتادة كنتيجة لحالتك الصحية النفسية. (مثلا الشعور بالاكتئاب او القلق) **خلال الاسابيع الأربعة الماضية**، هل تسببت **حالتك الصحية النفسية في:**	(اختر اجابة واحدة وضع علامة ✓ تحت الاجابة المناسبة) نعم	لا
أ) **التقليل من الوقت** الذي تقضيه في العمل او اي انشطة اخرى؟	☐	☐
ب) **التقليل** مما تود **انجازه** من العمل أو أي أنشطة أخرى؟	☐	☐
ج) **عدم انجاز** العمل او اي انشطة اخرى **بالحرص المعتاد**؟	☐	☐

الصحة الجسمية او النفسية

٦- **خلال الاسابيع الاربعة الماضية**، الى اي مدى تعارضت صحتك **الجسمية او النفسية** مع تأديتك لنشاطاتك الاجتماعية المعتادة مع عائلتك او اصدقائك او جيرانك او اي من المناسبات الاجتماعية الأخرى؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

- ☐ لم يكن هناك أي تعارض اطلاقا
- ☐ كان هناك تعارض قليل
- ☐ كان هناك تعارض متوسط
- ☐ كان هناك تعارض كبير
- ☐ كان هناك تعارض كبير جدا

شدة الألم

٧- ما شدة **الألم الجسمي** الذي عانيت منه خلال **الاسابيع الاربعة الماضية**؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

- ☐ لم يكن هناك أي ألم
- ☐ كان هناك ألم خفيف جدا
- ☐ كان هناك ألم خفيف
- ☐ كان هناك ألم متوسط
- ☐ كان هناك ألم شديد
- ☐ كان هناك ألم شديد جدا

٨- **خلال الاسابيع الاربعة الماضية**، الى اي مدى ادى **الألم الجسمي** الى التعارض مع تأديتك لأعمالك المعتادة (سواء داخل المنزل او خارجه)؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

- ☐ لم يكن هناك أي تعارض
- ☐ كان هناك تعارض قليل جدا
- ☐ كان هناك تعارض متوسط
- ☐ كان هناك تعارض كبير
- ☐ كان هناك تعارض كبير جدا

٩- الأسئلة التالية تتعلق بكيفية شعورك وطبيعة سير الأمور معك **خلال الأسابيع الأربعة الماضية**، الرجاء اعطاء اجابة واحدة لكل سؤال بحيث تكون هذه الاجابة هي الأقرب الى الحالة التي كنت تشعر بها.

خلال الأسابيع الأربعة الماضية، كم من الوقت:

(اختر اجابة واحدة وضع علامة ✓ تحت الاجابة المناسبة)

	في كل الأوقات	في معظم الأوقات	في كثير من الأوقات	في بعض الأوقات	في قليل من الأوقات	لم اشعر في أي وقت من الأوقات
أ) شعرت بأنك ملئ بالحيوية والنشاط؟	☐	☐	☐	☐	☐	☐
ب) كنت شخصا عصبيا جدا؟	☐	☐	☐	☐	☐	☐
ج) شعرت بأنك في حالة اكتئاب الى درجة لم يمكن معها ادخال السرور اليك؟	☐	☐	☐	☐	☐	☐
د) شعرت بالهدوء والطمأنينة؟	☐	☐	☐	☐	☐	☐
هـ) كانت لديك طاقة كبيرة؟	☐	☐	☐	☐	☐	☐
و) شعرت بالاحباط واليأس؟	☐	☐	☐	☐	☐	☐
ز) شعرت بأنك منهك (استُنْفِذت قُواك)؟	☐	☐	☐	☐	☐	☐
ح) شعرت بأنك شخص سعيد؟	☐	☐	☐	☐	☐	☐
ط) شعرت بأنك تعبان؟	☐	☐	☐	☐	☐	☐

١٠- **خلال الاسابيع الأربعة الماضية**، ما مقدار الوقت الذي تعارضت فيه **صحتك الجسمية او مشاكلك النفسية** مع نشاطاتك الاجتماعية (مثل زيارة الأصدقاء والأقارب وغير ذلك) ؟

(اختر اجابة واحدة وضع علامة ✓ أمام الاجابة المناسبة)

- ☐ كان التعارض في كل الأوقات
- ☐ كان التعارض في معظم الأوقات
- ☐ كان التعارض في بعض الأوقات
- ☐ كان التعارض في قليل من الأوقات
- ☐ لم يكن هنالك تعارض في أي وقت من الأوقات

١١- ما مدى **صحة او خطأ كل من** العبارات التالية (أ ، ب ، ج ، د) بالنسبة الى حالتك الصحية؟

(اختر اجابة واحدة وضع علامة ✓ تحت الاجابة المناسبة)

	صحيحة بلا شك	صحيحة غالبا	لا اعلم	خطأ غالبا	خطأ بلا شك
أ) يبدو أنني أصاب بالمرض أسهل من الآخرين.	☐	☐	☐	☐	☐
ب) حالتي الصحية مساوية لأي شخص أعرفه.	☐	☐	☐	☐	☐
ج) أتوقع أن تسوء حالتي الصحية.	☐	☐	☐	☐	☐
د) حالتي الصحية ممتازة.	☐	☐	☐	☐	☐

****** شكرا لتعاونكم ******