Reviewer’s report

Title: Normative Data and Discriminative Properties Of Short Form 36 (Sf-36) in Turkish Population

Version: 2 Date: 14 June 2006

Reviewer: Gemma Vilagut

Reviewer’s report:

This manuscript adresses metric characteristics of the turkish version of the SF-36 Health survey, one of the most widely used health-related quality of life instruments, in terms of data psychometric characteristics, discrimination of known groups (i.e. construct validity). It also provides normative data from a Turkish population is provided to facilitate interpretation of the scales scores.

Even though the SF-36 has been widely used and their scaling assumptions and metric properties have been evaluated in several translated versions, it is still important to evaluate this properties in each translated language to determine the extent to which equivalence of metrics with the original version has been achieved.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Page 4. It seems that a two-stage sampling design has been used. In the first stage, households were sampled by means of systematic sampling, and in the second stage, all individuals over 18 years within the household were selected to be interviewed. Although Systematic sampling method is more efficient than random sampling, it does not completely ensure that the final sample obtained from this type of survey will be representative of the general population. Also, the structure of clustered data (within household) may affect the generalizability of the results because members of the same cluster will be likely to be similar. It is specially relevant to ensure representativity of the sample in this case because one of the objectives of the manuscript was to obtain population norms, and norm-based comparisons require valid norms from a well defined and representative sample of the population of interest\[1\]. Therefore, the reference population should be appropriately defined, and representativeness of the data with respect to this reference population should be ensured and clearly justified. Is the sample distribution according to the main demographic variables similar to the population distribution? If necessary, weights should be applied to reflect the distribution of the general population.


Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Given that currently two different versions of the SF-36 are available, the authors should state that the version evaluated in this manuscript is SF-36 version 1.0.

Page 3, first paragraph, and page 8, second paragraph. The term "Healthy general population" may be misleading given the fact that general population is not free of disorders and, usually, individuals with disorders are not excluded from the calculation of general population norms.

Page 4, last paragraph. Mental ¤œfunctioning (MH) should be replaced by Mental ¤œHealth (MH).

Page 4, last paragraph. When the authors indicate the scale range, they should also provide the direction of the scores, i.e. meaning of the minimum and maximum values: 0 (worst) to 100 (best)

Page 5, first paragraph: Which normative group was used to obtain the mean and SD required to calculate
the z-scores standardizations of the scales? Were they US population data or Turkish data? When the country specific scales are obtained, it would be appropriate to use the country specific mean and SD in the z-scores transformation, but it is not clear which have been used in the article [1].

Page 5, paragraph 1: Although stated afterwards, it should be indicated here that the Turkish sample used to obtain the country specific factor weights was the same sample used throughout the whole manuscript.

Page 6, paragraph 1, RR for eligible people: Does the response rate of eligible people only take into account the households contacted? The person-level response rates should be computed by estimating the number of 18 or over year olds missed in nonrespondent households. For example, if households are selected with equal probabilities, RR should be based on counts of persons 18 or over sampled in respondent and nonrespondent households and not only taking into account respondent households. See AAPOR report for more detail [2]

Page 6. Paragraph 4. The authors state that Cronbach’s alpha= 0.90 is the generally accepted value for group comparisons, but it is actually the accepted value for individual comparisons. The accepted value for group comparisons is 0.7 [3]. Also, given the Cronbach’s alpha coefficients presented in table 2, RE scale also had internal consistency above 0.9 (0.93).

Page 6. Paragraph 4. PR should be changed by RP, as the term stands for Role Physical scale.

Page 6. Paragraph 4. Where it says PCS values ranged from 1.1, it should say from -1.1 (minimum value was found to be negative and the negative sign is missing in the text).

Page 6. Last paragraph: In the evaluation of construct validity of the SF-36, developers should list the hypotheses about specific logical relationships among relevant concepts or constructs at the end of the introduction, by specifying the domains of health status that would be expected to be more closely related to a criterion and the direction of the expected relationship.

Page 6. How do you explain the fact that women did not report poorer health in mental health related scales (MH and MCS)? Usually, women report worse scores than men in all scales, including those related to mental health.

Tables 3,4,6: It should be indicated which scoring method (country specific or standard algorithms) was used in the results given in tables 3,4 and 6.


