Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology

ISSN: 1017-7833 (Print) 1302-9657 (Online) Journal homepage: https://www.tandfonline.com/loi/tbcp20

Comparative Validity and Reliability Study of The QIDS-SR$_{16}$ in Turkish and American College Student Samples

Haluk Mergen (Associate Professor of Family Medicine), Ira H. Bernstein (Professor of Clinical Sciences), Vedide Tavli (Professor of Paediatric Cardiology), Kurtulus Ongel (Associate Professor of Family Medicine), Talat Tavli (Professor of Cardiology) & Seref Tan (Associate Professor of Faculty of Education)

To cite this article: Haluk Mergen (Associate Professor of Family Medicine), Ira H. Bernstein (Professor of Clinical Sciences), Vedide Tavli (Professor of Paediatric Cardiology), Kurtulus Ongel (Associate Professor of Family Medicine), Talat Tavli (Professor of Cardiology) & Seref Tan (Associate Professor of Faculty of Education) (2011) Comparative Validity and Reliability Study of The QIDS-SR$_{16}$ in Turkish and American College Student Samples, Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology, 21:4, 289-301, DOI: 10.5455/bcp.20110223124825

To link to this article: https://doi.org/10.5455/bcp.20110223124825

Published online: 08 Nov 2016.

DOI: 10.5455/BCP.20110223124825

Comparative Validity and Reliability Study of The QIDS-SR_{16} in Turkish and American College Student Samples

Haluk Mergen[1], Ira H. Bernstein[2], Vedide Tavli[3], Kurtulus Ongel[4], Talat Tavli[5], Seref Tan[6]

ÖZET:

Türk ve Amerikalı üniversite öğrencilerinde hızlı depresif belirti envanteri-özbildirim formu'nun ($HDBE_{16}$-ÖF) karşılaştırmalı olarak geçerlik ve güvenirliği

Amaç: Öğrenci ağırlıklı Türk örneklemine uygulanan Türkçe'ye çevrilmiş 16 maddelik Hızlı Depresif Belirti Envanteri-Özbildirim Formu'nun (HDBE16-ÖF): a) Amerikalı üniversite öğrencilerine uygulanan orijinal Amerikan versiyonu (QIDS-SR_{16}-US) ve b) aynı Türk öğrenci örnekleminde Beck Depresyon Envanteri-II (BDI-II) ile karşılaştırılarak geçerlik ve güvenirliğinin ortaya konması amaçlanmıştır. Çalışmamız Türkiye ve Amerika Birleşik Devletleri arasında yapılan bir kültürlerarası geçerlilik çalışmasıdır.

Metod: Uludağ Üniversitesi yerleşkesi Aile Sağlığı Merkezi'ne ayaktan başvuran öğrenci ağırlıklı 626 hastaya; www.ids-qids.org adresinden ulaşılabilen ve kısmen modifiye edilerek Türkçe'ye çevrilen $HDBE_{16}$-ÖF ve BDI-II testleri uygulandı. Ayrıca Güneybatı Teksas Üniversitesi'nde $HDBE_{16}$-ÖF envanterinin İngilizce orijinal versiyonu olan QIDS-SR_{16}-US 584 öğrenciye uygulanmıştır. Betimleyici istatistik, klasik açıklayıcı faktör analizi ve madde tepki kuramı analizleri, SAS ve MPlus istatistik programları ile yapılmıştır.

Bulgular: Türk deneklerin ortalama yaşı 21,1±2,16 (standart sapma) olup %67,8'i kadındı. Türk öğrencilerin aile içi depresyon öyküsü: annede %29, babada %8, kardeşte %14, kendisinde %16 ve akrabada %5 olarak bulundu. Amerikalı deneklerden 225 olguda hiç yaş belirtilmemiş haldeyken ortalama yaş 20.0±3,5 (standart sapma) ve tüm deneklerin %63,6'sı kadın olarak saptandı. $HDBE_{16}$-ÖF'nın madde ortalaması 6,94±4,85 (standart sapma) bulundu. $HDBE_{16}$-ÖF'ün iç tutarlılık katsayısı (Cronbach α) 0,78 idi ve ortalama madde-toplam korrelasyon katayısı 0,47 (0,33-0,61) bulundu. QIDS-SR_{16}-US'nin kaşılaştırılabilir madde ortalaması 6,09±3,76, Cronbach α 0,74, madde-toplam korrelasyon katsayısı 0,43 (0,24-0,54) olarak bulundu.

Hem $HDBE_{16}$-ÖF hem de QIDS-SR_{16}-US tek boyutlu iken BDI-II tek boyutlu olarak bulunmadı. $HDBE_{16}$-ÖF'ün ve QIDS-SR_{16}-US'un madde-total korelasyon ortalaması birbirine benzerdi. BDI-II ile $HDBE_{16}$-ÖF arasındaki korelasyon katsayısı 0.72 bulundu, bu değer disattenüe edildiğinde 0.90'a çıkmaktaydı. Çoklu grup doğrulayıcı faktör analizi $HDBE_{16}$-ÖF ve QIDS-SR_{16}-US'un aynı faktör yüküne sahip olduğu farklı değişik eşiklerinin olduğunu ortaya çıkarmıştır. Bu durum depresyon düzeyinde grup farklılıklarını ortaya koymaktadır. Türk deneklerin, Amerikalı deneklerden farklı olarak daha fazla depresyon geçirdikleri söylenebilir. Ayrıca $HDBE_{16}$-ÖF ve BDI-II'nın skorları birbirlerine eşitlenmiştir.

Tartışma: $HDBE_{16}$-ÖF'ün, hem Türkiye hem de Amerika Birleşik Devletleri'nde depresyon tanısı için çok sık kullanılan BDI-II testi gibi iyi psikometrik özellikleri ve yapısal geçerliliği olduğu saptanmıştır. Pek çok ortamda $HDBE_{16}$-ÖF'ün kullanılmasını önermekteyiz.

Anahtar sözcükler: Majör depresif epizod, tarama testi, geçerlilik, güvenilirlik

Klinik Psikofarmakoloji Bülteni 2011;21(4):289-301

ABSTRACT:

Comparative validity and reliability study of the QIDS-SR_{16} in Turkish and American college student samples

Objective: To evaluate the validity and reliability of the Quick Inventory of Depressive Symptomatology, self-reported version, in a Turkish student sample (QIDS-SR_{16}-T) by comparing it to (a) the American version (QIDS-SR_{16}-US) and (b) the Turkish version of the Beck Depression Inventory (BDI-II-T).

Materials and Methods: Slightly modified versions of the QIDS-SR_{16}-T, and the BDI-II-T were administered to 626 outpatients at the Uludağ University campus-based family health center. The QIDS-SR_{16}-US was administered to 584 respondents at an American university. SAS and MPlus were used to provide descriptive statistics, classical exploratory factor analysis, and item response theory analyses (in the form of a multiple group confirmatory factor analysis).

Results: The internal consistency (Cronbach α) of the QIDS-SR_{16}-T was 0.77. Both QIDS-SR_{16} versions were unidimensional, but the BDI-II-T was not. The mean QIDS-SR_{16}-T and QIDS-SR_{16}-US item-total correlations were similar. The correlation between the QIDS-SR_{16}-T and BDI-II-T was 0.72 (.90 when disattenuated). Multiple-group confirmatory factor analysis suggested that the QIDS-SR_{16}-T and QIDS-SR_{16}-US had the same factor loadings but different intercepts. This reflects group differences in level of depression, perhaps because the Turkish respondents, unlike their US counterparts, were seen in a medical context where illness-related depression is more prevalent. Scores on the QIDS-SR_{16}-T and the BDI-II-T were also equated.

Discussion: The QIDS-SR_{16}-T has good psychometric properties and convergent validity with the BDI-II-T. Its use is recommended when a self-reported instrument is appropriate.

Key words: Major depressive episode, screening scale, validity, reliability

Bulletin of Clinical Psychopharmacology 2011;21(4):289-301

[1]Associate Professor of Family Medicine, Uludağ University Family Health Center, Bursa-Turkey
[2]Professor of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, USA
[3]Professor of Paediatric Cardiology, Yeditepe University, İstanbul-Turkey
[4]Associate Professor of Family Medicine, Tepecik Research & Training Hospital, İzmir-Turkey
[5]Professor of Cardiology, Celal Bayar University, Manisa-Turkey
[6]Associate Professor of Faculty of Education, Uludağ University, Bursa-Turkey

Yazışma Adresi / Address reprint requests to: Haluk Mergen, M.D., A.S., Uludağ Universitesi Aile Sağlığı Merkezi, Görükle, Bursa-Turkey

Telefon / Phone: +90-532-441-9651

Elektronik posta adresi / E-mail address: haluk.mergen@gmail.com

Gönderme tarihi / Date of submission: 01 Şubat 2011 / February 01, 2011

Kabul tarihi / Date of acceptance: 23 Şubat 2011 / February 23, 2011

Bağıntı beyanı:
H.M., I.H.B., V.T., K.Ö., T.T., S.T.: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemişlerdir.

Declaration of interest:
H.M., I.H.B., V.T., K.Ö., T.T., S.T.: The authors reported noconflict of interest related to this article.

INTRODUCTION

A major depressive episode (MDE) must last at least two weeks and involve five of nine core symptoms: (a) sleep disturbance, (b) sad mood, (c) change in appetite and/or weight, (d) difficulty in concentration and decision making, (e) negative self view, (f) thoughts of death or suicide, (g) loss of general interest, (h) reduced energy level, and (i) restlessness or agitation. One of the symptoms must be (b) or (g) (1). Estimates of lifetime MDE prevalence in different countries r ange from 5 to 17% with an average estimate of 12%. Various prevalence estimates obtained in Turkey include 8.4% (2), 26.2% (3), and 39.4% (4).

Because of the high prevalence of depression, accurate, time-efficient measurement of depressive symptom severity is of great importance in conducting cost-efficient clinical trials (5). Self-reports are useful to both clinicians and researchers who wish to monitor treatment outcomes in a time- and cost-effective manner.

One way to evaluate the translation of a test is to compare the results from that sample to one obtained from a sample that is fluent in the original language. In the present case, we had samples of Turkish and American college students. These samples are similar in terms of cognitive abilities, but as will be noted, are not similar in level of depression. Another aspect of the evaluation is to see how well the measure correlates with an accepted measure, both of which are available as translations, i.e., its convergent validity. In the present case, the Beck Depression Inventory-II (BDI-II) is one such instrument. This study evaluated the validity and reliability studies of the QIDS-SR_{16}-T by comparing our results obtained in our sample to the American version, the QIDS-SR_{16}-US. and by correlating the QIDS-SR_{16}-T version with the BDI-II-T.

METHODS

The 16 item self-reported version of the Quick Inventory of Depressive Symptomatology (QIDS-SR_{16}) was first described by Rush et al. (2003) (6) and also exists in clinician rated and interactive voice formats (6). The QIDS-SR_{16}-T was obtained from the www.ids-qids.org website. The translation and back-translation of the QIDS-SR_{16} was done by a translation team into 30 languages. A few minor changes were made to accommodate actual Turkish usage, e.g., references to weight in pounds was replaced by weight in kilograms. The specific translation of the QIDS-SR_{16}-T used in this study is available from the first author. The QIDS-SR_{16} requires minimal training to administer and is freely available for use (7). The 16 items span the nine core symptom domains as defined by the text revision of version IV of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) (8). These items are sleep disturbance (items 1-4), depressed (sad) mood (item 5), change in appetite or weight (items 6-9), concentration/ decision making (item 10), self view (item 11), suicidal ideation (item 12), interest (item 13), energy/fatigue (item 14), and psychomotor agitation/retardation (items 15 and 16) (8). The determination of remission or partial remission is based on the DSM-IV-TR, which recommends that all nine diagnostic criterion symptoms be assessed (5,8). The responses for each item range from zero to three, with zero indicating the absence of that symptom in the past week, and the remaining categories defining mild, moderate, and severe presence of the symptom in question in that time period. The scoring scheme involves adding the scores for the nine symptom domains to yield a total score that consequently ranges from 0 to 27. QIDS-SR_{16}-US scores are commonly interpreted as follows: no depression (0-5), mild depression (6-10), moderate depression (11-15), severe depression (16-20), and very severe depression (≥21) (6); however, the Turkish version of the QIDS-SR_{16}, QIDS-SR_{16}-T, has not been evaluated for its reliability and validity. The purpose of this study was to perform this evaluation.

The inventory used for comparison, was the Beck Depression Inventory-II, a revision of the original scale published in 1996 by Beck et al (9,10). The Turkish version, the Beck Depression Inventory-II-Turkish (BDI-II-T), was validated by Tegin (11). The main advantages of the QIDS-SR_{16} are its brevity and survey of the nine major psychiatric symptoms as listed in the DSM-IV-TR. In comparison, the BDI-II includes 21 items, the Hamilton depression scale has 17, 21, 24, 28, or 31 items and the Zung Self-Depression Scale has 20 items. The QIDS-SR_{16} also provides: 1) equivalent weightings (0-3) for each symptom item, 2) clearly stated anchors that estimate the frequency and severity of symptoms; and 3) matched clinician and patient ratings between the scale's clinician

and self-reported forms (6,12).

The Turkish sample consisted of students, who presented consecutively to the outpatient clinic of the university. On the other hand, the US sample was composed of students taking a course in introductory psychology at a university in the Southwestern United States. Since they did not present or seek help at a clinic, this sample was unselected with regard to psychological problems and thus probably was less depressed as a group than the Turkish sample. One case was eliminated from each sample due to errors in filling out the answer sheets.

The Turkish study conformed to the Helsinki declaration requirements and had approval from the University ethics committee. The Turkish students gave informed consent and voluntarily participated in the study. They completed the QIDS-SR_{16}-T and BDI-II-T in 7-10 minutes. Out of 670 patients' questionnaires, 44 questionnaires were eliminated because of problems such as failure to complete the two tests. The gathering of the 584 QIDS-SR_{16}-US questionnaires met similar ethical guidelines and was approved by the University Institutional Review Board. One US questionnaire was eliminated due to a coding error. Thus, the Turkish and US samples consisted of 626 and 583 respondents, respectively.

Statistical Analyses

Descriptive analyses, classical test theory, exploratory factor analysis and item response theory analyses were used. An item response analysis is a form of confirmatory factor analysis that tests proposed item structures. A single group analysis was performed in the case of BDI-T, and multiple group (Turkish vs. US) analyses in the case of the QIDS-SR_{16}. The descriptive analyses were evaluated using χ^2 tests for discrete variables and t-tests for continuous variables. The classical test theory analyses generated item means, item standard deviations, item-total correlations (r_{it}), scale means, and scale standard deviations. Exploratory factor analysis was used to evaluate the dimensionality of the three tests. This analysis is important for its own sake, but is also important for the item response analyses as well. Parallel analysis (13-16) was used to decide upon the number of factors (dimensionality). The version used in this study consisted of generating a series of 50 matrices of random normal deviates. Each matrix had 9 columns and 626 rows for the QIDS-SR_{16}-T, 9 columns and 583 rows for the QIDS-SR_{16}-US, and 21 columns and 626 rows for the BDI-T. Thus, the number of columns equaled the number of items on the actual test and the number of rows equaled the number of participants in the relevant sample. Each of the resulting 150 matrices was then subjected to principal component analysis and the eigenvalues averaged within each of the three sets. A test was considered unidimensional if (a) its first eigenvalue exceeded the average of the simulated eigenvalues for that test and (b) all subsequent obtained eigenvalues were smaller than the corresponding simulated eigenvalues. Because the essential features of the exploratory analysis were also present in the item response theory analysis and the latter will be discussed starting in the next paragraph, the former will not be presented except to note that loadings on the first (only) factor ranged from .45 (item 6, suicidal ideation) to .74 (item 2, sad mood). All of these analyses used SAS 9.2

Because we were not comparing the Turkish sample with that of another nationality, the item response theory analysis for the BDI-II-T follows the standard two parameter logistic polytomous format (17,18). Thus, the probability of an affirmative response is given by $P(\theta/_{ij}) = \frac{1}{1 + e^{a_i(\theta - b_{ij})}}$. In this equation, a_i describes the 21 item slopes (factor loadings) and b_{ij} describes the 63 (21 items x 3 dichotomous criteria) intercepts (locations) of the functions. These dichotomies are 0 vs. 1, 2, or 3 (no pathology vs. at least some pathology), 0 or 1 vs. 2 or 3 (no or mild pathology vs. moderate or severe pathology), and 0, 1, or 2 vs. 3 (no pathology, mild pathology, or moderate pathology vs. severe pathology). Finally, θ denotes the magnitude of the latent variable (depression in this case). In order to make the equations estimable, the value of a_i is the same for all three intercepts so they form a set of three (in the present case) parallel functions. The larger a_i is, the steeper the slope of the functions, and the more discriminating they are with regard to θ. Though it comes from a different theoretical perspective, a_i serves a similar role to the r_{it} of classical test theory, and the values of a_i and r_{it} are typically correlated over items. The b_{ij} describes the tendency to choose the higher category of each dichotomy. The higher its value, the more likely the lower alternative is chosen and, if at 0, the intercept is said to be at threshold, as the probabilities of the two alternatives are each .5. The values of b_{ij} serve a role similar to that of the item means in classical test theory, but there are three such

values/item in the present cast per set of functions vs. item one mean. Both a_i and b_{ij} are scaled in a z-score metric. The θ also has a factor variance, which was set at 1.0 to define the unit of measurement.

The two versions of the QIDS-SR_{16} generate a series of multiple group models, which evaluate differences between the Turkish and US samples, thus complicating the basic Samejima model. The series incorporate various nested comparisons as described below. Two models are nested when one is made more specific than the other in a particular way. In the present case, this meant comparing the fit of one model where certain parameters in the Turkish and US model were allowed to differ vs. a model in which they were required to be the same. The difference in fit between the two models can be tested to see if the equality constraint is tenable vs. whether the two groups require separate sets of parameters. There are a total of five models, described more fully in the next paragraph as models 1, 2a, 2b, 3, and 4. Each model generates five sets of estimates/group. By analogy to the BDI-II-T, there are 9 slopes which describe the strengths of relation between the nine domains (symptoms) and (θ) overall depression. Following standard terminology, these are denoted a_i and there are 27 locations which describe the tendencies to choose the more pathological of the particular dichotomy. These are denoted b_{ij} where *i* denoted the 9 domains and *j* denoted the 3 dichotomies per domain as described with the analysis of the BDI. There is also a group mean for θ for each nationality, which was set at 0 in all but one of the nested comparisons to be presented. Fourth, there is a variance of θ for each nationality, which was always 1.0. Finally, there is a residual variance associated with each of the nine QIDS-SR_{16} domains. These factor variances were all fixed at 1.0, and the residual variances were all fixed at 1.0, though the results of additional analyses, not reported, allowing the factor variances to differ from 1.0, did not differ materially from those to be presented.

The first of the series of QIDS-SR_{16} models, model 1, allowed the values of a_i and b_{ij} to differ for the Turkish and US groups, i.e., the models were fit separately to the two nationalities. Four additional models introduced various constraints. Model 2a constrained a_i to equality between groups while letting values of b_{ij} vary freely. Thus, the model assumed equal slopes (discriminations, factor loadings) in the Turkish and US samples. Conversely, model 2b constrained b_{ij} to equality within groups while letting a_i vary freely, so it assumed equal intercepts in the two samples. Model 3 constrained both a_i and b_{ij} to equality, thus assuming both equal slopes and equal intercepts, but allowed the group means to differ. This assumes that any difference between nationalities is proportional to the factor loadings and also assumes equal locations. Finally, model 4 constrained a_i and b_{ij} to equality while constraining the group means to 0, i.e., effectively treating the groups as identical. Model 1 was evaluated by testing its overall fit in terms of χ^2 whereas the remaining models were evaluated by means of the Bentler-Satorra χ^2 difference test (19). The latter is needed because weighted least squares estimation was used rather than maximum likelihood estimation, given the categorical nature of the responses, so simple differences in χ^2 could not be used to test the national differences. Models 2a, 2b, and 3 were tested against model 1, and model 4 was tested against model 3. However, even well-fitting models may generate a significant χ^2 because of incidental issues like nonnormality. As a result, descriptive measures of fit have assumed greater importance in evaluating models. In particular, it is common to require the confirmatory fit index (CFI) to exceed .95 and the root mean square error of approximation (RMSEA) to be less than .05, and this strategy was emphasized in this study. In general, if a particular model is unacceptable, further constraints on that model are also unacceptable. As a results, if model 3 can be rejected, so can model 4, though all results will be presented for completeness.

The chosen model also provides a test information function (TIF) for each of the three tests. This describes the sensitivity of θ to change as a function of its level. The TIF serves a similar purpose to coefficient α in classical test theory but is a function of θ instead of a constant. Finally, we used the procedure described in (20) to equate the QIDS-SR_{16}-T to the BDI-II-T. This involves finding the expected a posteriori value of θ for each raw score. Scores on the two tests are considered equated if they have the same or highly similar values of θ. This assumption is met because the same participants took both tests. Mplus was used for these item response theory analyses.

RESULTS

The mean age of the Turkish participants was 21.1±2.16 years and 67.8% were female. The demographic data of

Table 1: Sociodemographic and medical properties of the Turkish subjects.

	N	%
Age		
<20 years old	253	40.3
20-24 years old	354	56.4
>24 years old	21	3.3
Education		
Illiterate	6	1
Primary school	3	0.5
High school	18	2.9
University	601	95.7
Income		
<500TL	243	38.7
501-1000TL	218	34.7
1001-2000TL	138	22.0
>2000TL	29	4.6
Work		
Does not work	523	83.3
Does work	105	16.7
Dwelling		
Village	86	13.7
Town	88	14
City	454	72.3
Number of sibling		
No sibling	66	10.5
1 sibling	194	30.9
2 siblings	135	21.5
3 siblings	122	19.4
>3 siblings	111	17.7
Family Depression History		
N/A	347	55.3
No	224	35.7
Yes	57	9.1

Note: percentages for a given variable may not add to exactly 100.0 because of rounding error.

outpatients are illustrated in Table 1. Some of the students had a family history of depression as well: mother 29%, father 8%, sibling 14%, self 16%, and relative 5%. The mean age of the American subjects was 20.0±3.5 years with 225 missing observations, and 63.6% were female. No other demographic data were available. The Turkish subjects were slightly, but significantly older than the American subjects, $t(987) = 4.35$, $p < .0001$. However, the percentages of males and females were the same in the two ethnic groups, $\chi^2(1) = 1.84$, $p = 0.17$.

Both QIDS-SR_{16}-T and BDI-II-T scores were higher in women than men (t=2.97, p=0.03 and t=2.61, p=0.009). QIDS-SR_{16}-T scores related to a family history of depression (t=2.08, p=0.038). The incidence of depressive symptomatology was significantly greater among females than males in six domains: (a) sleep disorders, (b) weight problems, (c) irritability, (d) sadness, (e) concentration/ decision making, and (f) energy level. The t values for these gender differences ranged from 2.09 to 4.79, and the associated p values ranged from .023 to .0001. In addition, greater irritability was associated with higher monthly income, t= 2.05, p= .041. Thoughts of death or suicide were more frequent in students with a family history of depression, t= 2.86, p= .004 and in older students, t= 2.34, p= .019. Low energy level was also more common among those with a family history of depression, t= 2.70, p= .007.

CTT Analysis

The scale mean (sd) of the QIDS-SR_{16}-T was 6.94±4.85. The internal consistency reliability (Cronbach's α) was 0.78, and the mean r_{it} was 0.47 with a range of r_{it} values from 0.33 to 0.61. The comparable mean (sd) of the QIDS-SR_{16}-US was 6.09±3.76. Cronbach's α was .74, the mean r_{it} was .43, and the range of r_{it} was from .24 to 54.

Table 2 contains the QIDS-SR_{16}-T and QIDS-SR_{16}-US item statistics. The QIDS-SR_{16}-T mean was significantly higher than the QIDS-SR_{16}-US mean, t= 3.35, p< .0001. Table 3 contains the comparable BDI-II-T statistics. Within the Turkish sample, the correlation between the QIDS-SR_{16} and the BDI21 was .75. Disattenuating by dividing this obtained correlation by the square root of the product of the alpha coefficients (21) indicates that the correlation between the underlying traits measured by the two scales is .90. Thus, they are highly similar, but not identical, measures of depressive symptomatology.

Table 2: QIDS-SR16 item and scale statistics: item means, item standard deviations, item-total correlations (r_{it}), sample sizes (N), raw Cronbach's α, scale means and scale standard deviations for Turkish and US samples

	Turkey			US		
Domain	Mean	Std	r_{it}	Mean	Std	r_{it}
1	1.65	.94	.37	1.65	.80	.24
2	.77	.89	.61	.63	.68	.54
3	.86	.93	.39	.91	.88	.34
4	.92	.86	.61	.57	.70	.49
5	.76	1.23	.49	.43	.80	.39
6	.13	.40	.33	.23	.54	.41
7	.43	.74	.44	.38	.67	.45
8	.51	.73	.51	.55	.69	.49
9	.89	1.12	.52	.75	.75	.52
N	626			583		
Raw α	.78			.74		
Scale Mean	6.94			6.09		
Scale Std	4.85			3.76		

Table 3: BDI-II-T item and scale statistics: item means, item standard deviations, item-total correlations (r_{it}), sample size (N), raw Cronbach's α, scale mean and scale standard deviation

Item	Mean	Std	r_{it}
1	.40	.68	.63
2	.34	.53	.49
3	.26	.60	.44
4	.50	.60	.62
5	.60	.61	.60
6	.42	.73	.50
7	.33	.59	.65
8	.49	.67	.57
9	.09	.38	.37
10	.50	.97	.45
11	.55	.80	.49
12	.46	.71	.62
13	.59	.78	.65
14	.23	.63	.45
15	.54	.64	.55
16	.47	.68	.49
17	.33	.64	.51
18	.24	.52	.42
19	.09	.33	.23
20	.28	.57	.32
21	.34	.73	.37
N	620		
Raw α	.89		
St. Alpha	.89		
Scale Mean	8.03		
Scale Std	7.61		

Exploratory Factor Analysis (Dimensionality)

The QIDS-SR_{16}-T and the QIDS-SR_{16}-US were both unidimensional by the parallel analysis criterion. Specifically, the first eigenvalues of these two tests were 3.39 and 3.11 vs. 1.19 and 1.19 for the simulated data, and the second eigenvalues were .92 and 1.02 vs. 1.13 and 1.13 for the simulated data. Additional obtained eigenvalues were all less than their simulated counterparts. However, the BDI-II-T was two dimensional as its first three eigenvalues were 6.79, 1.46, and 1.10 vs. 1.34, 1.28, and 1.24 for the simulated data (all subsequent real eigenvalues were less than their simulated counterparts). Following a promax rotation, the two Beck factors correlated .42. Equally important, the first factor accounted for considerably more variance (22%) than the second (8%), so the latter is fairly minor. This second factor is also difficult to interpret as was defined by items number 16 (sleep problems), 18 (appetite), 19 (weight loss), 20 (hypochondriasis), and 21 (interest in sexuality). These results indicate the assumption that a scale measures only one dimension, which is relevant to the item response analysis of the next section, is met for both versions of the QIDS-SR_{16}, but the Beck analysis should be interpreted with some caution because of the presence of a minor second factor.

Main IRT Analyses

Table 4 contains the fit statistics for the various multiple group models. The model χ^2 for the baseline model 1 testing the two forms of the QIDS-SR_{16}, which fit parameters separately to the two groups, was 133.68 and is significant beyond the .0001 level. However, the CFI for model 1 was .981 and the RMSEA was .049 so the fit of this baseline model is acceptable in descriptive terms. The difference χ^2 comparing models 1 and 2a was a significant 19.56 on 9 df, $p < .05$, but the CFI and RMSEA were identical to that observed with model 1. Conversely, the χ^2 comparing models 1 and 2b was 208.44 on 27 df, and the CFI and RMSEA were .926 and .079. Model 2b can be rejected by these criteria. Accordingly model 2a was tentatively accepted, i.e., it was assumed that QIDS-SR_{16}-T and QIDS-SR_{16}-US have the same slopes (relations of θ to the nine domains) but different thresholds (levels) in the two groups. The poor fit of model 2b makes testing of model 3 irrelevant which, in turn, also makes testing model 4 irrelevant.

Table 4: Fits of the multiple group models

Model	Intercepts	Slopes	Means	χ^2	df	p	RMSEA	CFI
1	Free	Free	Constrained	133.68	54	.00	.05	.98
2a	Free	Constrained	Constrained	19.56	9	.02	.04	.98
2b	Constrained	Free	Constrained	208.44	27	.00	.08	.93
3	Constrained	Constrained	Free	24.97	35	.00	.08	.92
4	Constrained	Constrained	Constrained	14.	1	.00	.08	.91

Note: RMSEA – root-mean square error of approximation, and CFI = comparative fit index. The χ^2 for model 1 tests the overall fit of the model whereas the χ^2 for the remaining models tests the difference in fit (model 1 is used to test models 2a, 2b, and3, and model 3 is used to test model 2). As a form of weighted least squares was used in testing, the Satorra-Bentler adjustment was employed in model testing.

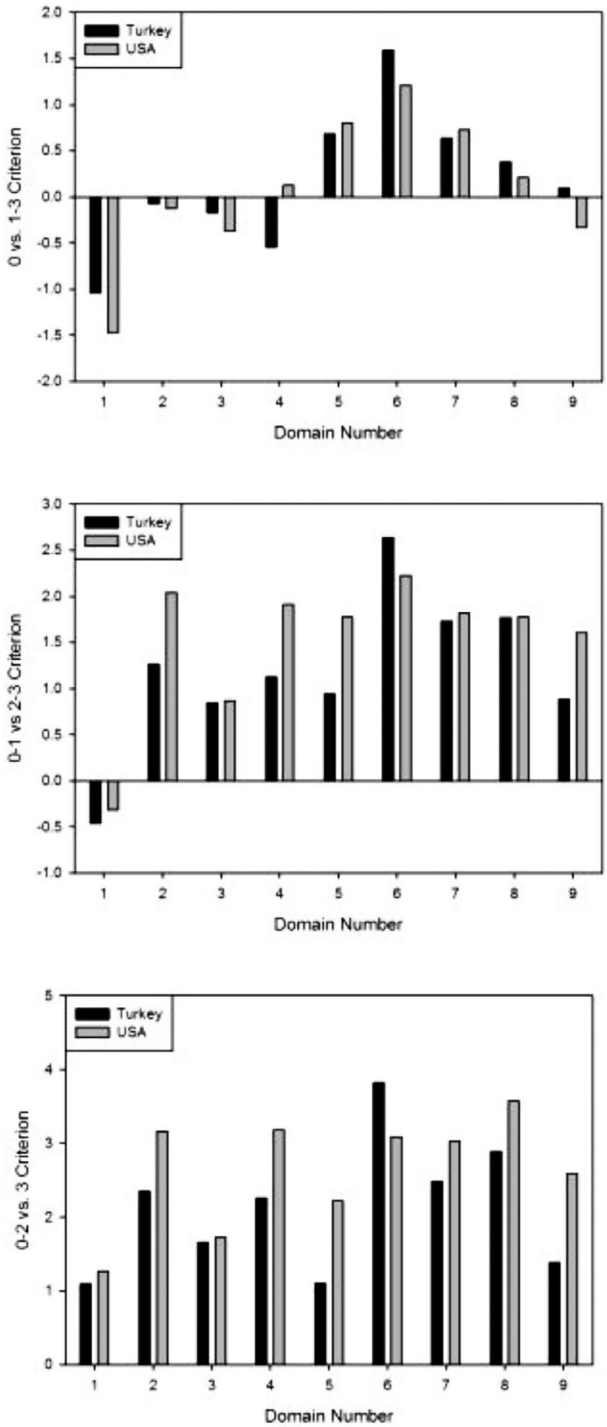

Figure 1: Domain thresholds for the three criteria (0 vs. 1-3 in the top panel, 0 and 1 vs. 2 and 3 in the middle panel, and 0-2 vs. 3 in the bottom panel), separately for the Turkish and US groups.

Figure 1 contains the thresholds for the three criteria (0 vs. 1-3 in the top panel, 0-1 vs. 2- 3 in the middle panel, and 0-2 vs. 3 in the bottom panel), separately for the Turkish and US groups. Note that contrary to what was obtained in the classical analysis of the total tests scores, QIDS-SR_{16}-US 0 vs., 1-2 thresholds were actually lower than QIDS-SR_{16}-T 0 vs., 1-3 thresholds, implying that American respondents were more willing to report a pathological category than Turkish respondents. However, the remaining two comparisons were in line with the classical test theory analysis; QIDS-SR_{16}-T thresholds were lower than QIDS-SR_{16}-US thresholds 8 of 9 times in each case, implying the Turkish participants were the more willing to endorse the moderate and severe categories of depressive pathology..

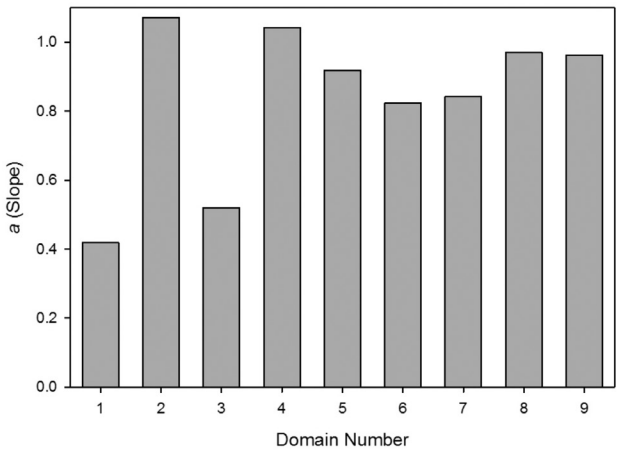

Figure 2: Common domain slopes

Figure 2 contains the common values of the slope. As is usually the case, these slopes parallel the item-total correlations of the classical test analysis. Domain 2 (sad mood) has the highest value (is most discriminating) and domains 1 and 3 the lowest (are least discriminating). The remaining domains are closer to domain 2 than domains 1 and 3. Figures 1 and 2 thus describe the main features of the QIDS-SR_{16} item structures. The BDI's structure is of lesser interest and does not meet the item response theory model's assumption that a single dimension underlies the data so it will not be presented. However, it is available upon request from the authors.

Test Information Functions

Figure 3 contains the test information functions (TIF) for the Turkish QIDS-SR_{16}, the QIDS-SR_{16}-US and the BDI-II-T. As noted, these represent the change in the respective measure per change in depression (θ generically).

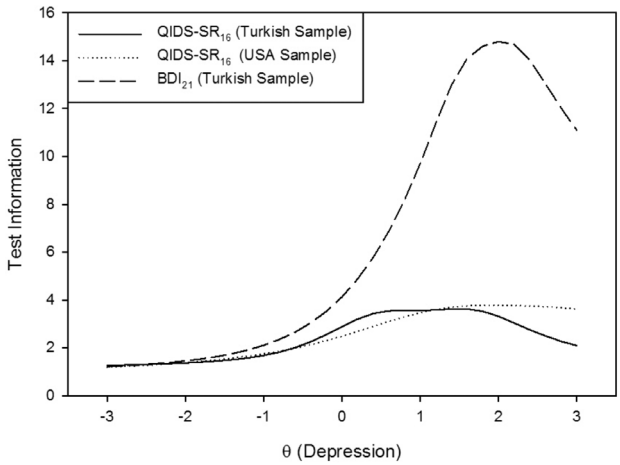

Figure 3: Test information functions (TIF) for the Turkish and US versions of the QIDS-SR_{16} and for the BDI-II-T

The points to note here are: (a) the high degree of similarity between the two versions of the QIDS-SR_{16}, (b) the fairly substantial similarity among all three measures below the latent variable mean (θ= 0), (c) the tendency of the BDI-II-T to be most discriminating past this point, and (d) the fall-off for the BDI-II-T at the positive tail (θ > 2).

Test Equating

Table 5 contains the results of equating the QIDS-SR_{16}-T and BDI-II-T. Thus, a raw score of 8 on the QIDS-SR_{16}-T equates to a raw score of 9 on the BDI-II-T because both lead to an estimated θ of .7 on the normally distributed depression scale. Not all values equate exactly so expected a posteriori (EAP) values within ± .1 θ units were accepted as matching, e.g., a QIDS-SR_{16}-T score of 4 was treated as matched to a BDI-II-T score of 3 as they produced EAP values of -.4 and -.5 respectively.

Table 5: Results of equating the QIDS-SR_{16}-T to the BDI-II-T

QIDS-SR_{16}-T		BDI-II-T	
Score	**EAP**	**Score**	**EAP**
0	-1.4	0	-1.3
1	-1.1	1	-1.0
2	-.9	1	-1.0
3	-.6	2	-.7
4	-.4	3	-.5
5	-.2	4	-.3
6	.1	5	.0
		6	.2
7	.4	7	.4
		8	.6
8	.7	9	.7
		10	.9
9	1.0	11	1.0
		12	1.1
10	1.2	13	1.2
		14	1.4
11	1.5	15	1.5
		16	1.6
12	1.7	17	1.7
		18	1.8
13	1.9	19	1.9
		20	2.0
14	2.1	21	2.1
		22	2.1
		23	2.2
15	2.3	24	2.3
		25	2.4
16	2.5	26	2.5
17-18	2.6	27	2.6
19	2.7	28	2.7
20	2.8	29	2.8
21	2.9	30	2.9
22	3.0	31	3.0
23	3.1	32	3.1
24	3.2	33	3.2
25	3.3	34	3.3
		35	3.4
		36	3.5
		37	3.6
26	3.7	38	3.7
		39	3.7
		40	3.8
27	3.8	41	3.8
		42-45	3.9
		[3] 46	4.0

Note: EAP = Expected a posteriori value of θ (depression)

DISCUSSION

A major finding was the essential equivalence of the loadings (trace line slopes) of the domains in the two samples, meaning that each domain measured depression to the same extent in the two cases. These values are similar to those previously obtained from the QIDS-SR_{16}-US, e.g. (5, 22-29). Thus, the scale is unidimensional. Sad mood relates most strongly to overall depression and suicidal ideation relates least strongly, as noted in Fig. 2, and as reported in various studies conducted on US samples. The values of coefficient α reflect the sample variances, as is true in the various studies. This is in part due to the fact that we did not have many severely depressed patients in our sample whereas other studies tend to run the gamut of depression. The difference in intercepts (levels) is more difficult to interpret as it may reflect either the fact that the Turkish respondents were recruited from a medical setting where depression is perhaps more common and the

American respondents were not, or other sample characteristics or differences due to translation. At this point, the former seems the more reasonable alternative. The second major finding was the high correlation among Turkish respondents between the BDI-II-T and the QIDS-SR_{16}-T, which is strong evidence for the latter's convergent validity. However, one difference between the two is that the BDI-II-T contains at least two dimensions, whereas the QIDS-SR_{16}-T is unidimensional, which is perhaps responsible for the lack of perfect correlation. This difference may also help explain the difference in TIF past the group mean on θ, i.e. mean depression level.

There is one important point regarding our test equating. We treated the QIDS-SR_{16}-T and the BDI-II-T as two self-descriptive measures of depressive symptomatology with neither serving as a "gold standard" to define depression for the other. This led to equating the two tests in terms of common inferred (θ) values. This is quite different from the ROC approach used by Bilgel and Bayram (30) which does accept the BDI-II-T as a "gold standard". We have used ROC analysis in some earlier studies, e.g., (25) but the criterion to define depression was a structured clinical interview and not a self-report of symptomatology. We accept that the interview is closer to a "gold standard" than is another self-reported measure.

Lamoureux et al. (7) have provided another example of a study for which ROC analysis was appropriate. They studied 155 heterogeneous primary care outpatients, similar to our validity study. They used both the clinician-rated and self-reported QIDS scales, which they compared with the results of the Structured Clinical Interview for DSM Disorders (SCID). They reported an area under the curve of 0.82. The value of Cronbach's α was 0.86, which is somewhat greater than our value, perhaps reflecting differences in sample variability. They suggested a total score cutoff of 13-14 for moderate depression, which yields a sensitivity of 76.5% and specificity of 81.8%. They emphasized the need for screening of MDE in primary care, which could substantially improve patient outcomes, particularly when combined with efforts to promote adequate treatment and follow-up.

Several other studies illustrate the wide range of settings, in which the QIDS-SR_{16}-US has been applied. These are important in this context given the similarity of the two versions of the QIDS-SR_{16}. Bernstein, Rush, Carmody et al. (23) studied a low income and relatively low education public sector sample using, as here, both classical test theory and item response theory analyses. Overall, the self-reported and clinical versions of the QIDS were similar in their psychometric properties. Similarly, Rush et al. (5), found the clinical and self-reported versions of the QIDS to compare well to the Hamilton Rating Scale for Depression, which is perhaps the most widely used measure to evaluate depressive symptomatology in the United States. Likewise, Doraiswamy, Bernstein Rush, et al. (27) found the Montgomery-Åsberg Depression Rating Scale (MADRS), QIDS-C_{16}, and QIDS-SR_{16} to perform similarly in an elderly population, where the MADRS is perhaps the most widely used measure in Europe. The α coefficients ranged from 0.85 to 0.89. Moreover, Bernstein et al. (31) also found the QIDS scales effective in the evaluation of patients with bipolar disorder. In addition, all of these papers found the QIDS scales to be unidimensional. Other studies noted that QIDS scales are similar in reliability to comparable instruments used to evaluate depressive symptomatology (6,9,32-33).

Brown et al. studied the QIDS-SR_{16}, IDS-SR_{30}, $HRSD_{17}$ and Mini Asthma Quality of Life Questionnaire in asthmatic patients at treatment exit because of the highly co-occurrence of asthma with depression (34). Cronbach α values were highest (0.95) for the IDS-SR_{30}; because of its greater length. These values were 0.87 for the QIDS-SR_{16} and the $HRSD_{17}$. QIDS-SR_{16} and $HRSD_{17}$ total scores are highly correlated (r=0.85) as are QIDS-SR_{16} and IDS-SR_{30} scores (r=0.97). All three scales used in the Brown et al. study, showed comparable sensitivity to symptom change.

Bernstein, Rush, Yonkers et al. (23) had postpartum patients and non-postpartum female controls take the QIDS-SR_{16}-US. Both groups showed low energy level and restlessness/agitation. However, the non-postpartum group reported greater sad mood, suicidal ideation, and reduced interest. Conversely, the postpartum group exhibited psychomotor symptoms (restlessness/agitation) and impaired concentration/decision-making. Carmody et al. (35) compared the QIDS-SR_{16}-US to the Montgomery Åsberg Depression Rating Scale and used the Orlando et al. (31) procedure to equate the two sets of scores.

In summary, the QIDS-SR_{16}-US has been used successfully in a wide variety of settings, and the QIDS-SR_{16}-T appears sufficiently similar to suggest its use in a variety of settings to screen for depression, including primary care settings.

Limitations

Because the study sample primarily consists of university students seen at an outpatient clinical setting, it would be of use to apply it to a more general Turkish population, especially an academic one that is similar to the present American sample. A second important limitation is that we did not have repeated test data to evaluate stability and sensitivity to change of the Turkish QIDS-SR_{16}, which would be necessary if the scale were to monitor therapeutic effects. However, this was not the goal of this study. In addition, sensitivity to change was evaluated in the QIDS-SR_{16}-US and found at least adequate (36). Given the similarity of the two versions, this is important, albeit indirect evidence.

Although the details of the item responses analysis of the Beck are less important than those of the QIDS-SR_{16}, the statement, expressed earlier regarding the need to interpret the results with caution due to the presence of a minor second factor, is useful to note.

References:

1. Sadock BJ, Sadock VA. Kaplan and Sadock's Synopsis of Pscyhiatry: Behavioral Sciences/Clinical Psychiatry (10th edition), Philadelphia: Wolters Kluwer, Lippincott Williams&Wilkins; 2007. p.527-8.
2. Mergen H, Öngel K. Factors associated with depression among Turkish faculty of education freshmen by Beck Depression Inventory–II-Turkish [Dejavniki depresivnosti pri študentih prvega letnika pedagoške fakultete z vprašalnikom BDI-II-T]. Zdrav Vestn 2009; 78:548-54.
3. Bostancı M, Ozdel O, Oguzhanoglu NK, Ozdel L, Ergin A, Ergin N, et al. Depressive Symptomatology Among University Students in Denizli, Turkey: Prevalence and Sociodemographic Correlates. Croat Med J 2005; 46(1):96-100.
4. Mergen H, Erdoğmuş Mergen B, Tan Ş, Öngel K. Evaluating The Depression and Related Factors Among the Students of the Faculty of Education at Celal Bayar University. The New Journal of Medicine 2008; 25:169-174.
5. Rush AJ, Bernstein IH, Trivedi MH, Carmody TJ, Wisniewski S, Mundt JC, et al. An evaluation of the quick inventory of depressive symptomatology and the hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. Biol Psychiatry 2006; 59(6):493-501.
6. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biological Psychiatry 2003;54(5):573-83.
7. Lamoureux BE, Linardatos E, Fresco DM, Bartko D, Logue E, Milo L. Using the QIDS-SR16 to identify major depressive disorder in primary care medical patients. Behav Ther 2010; 41(3):423-31.
8. American Psychiatric Association. Diagnostic and Statistical Manual for Mental Disorders (DSM-IV), Fourth Edition. Washington, DC: American Psychiatric Association, 1994.
9. Beck AT, Steer RA, Brown GK. Beck Depression Inventory-Second Edition Manual. San Antonio, TX: The Psychological Corporation, 1996.
10. Beck AT, Ward C, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961; 4: 561-71.
11. Tegin B. Depresyonda bilişsel bozukluklar Beck modeline göre bir inceleme Yayınlanmamış doktora tezi. [Cognitive Malfunctions in Depression. Survey according to Beck Model, unpublished dissertation]. Hacettepe Üniversitesi, Psikoloji Bölümü; 1980.
12. Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, et al. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. Neuropsychopharmacology 2005; 30(2):405-16.
13. Horn JL. An empirical comparison of various methods for estimating common factor scores. Educ Psychol Meas1965; 25: 313-22.
14. Humphreys LG, Ilgen D. Note on a criterion for the number of common factors. Educ Psychol Meas1969; 29: 571-578.
15. Humphreys LG, Montanelli RGJr. An investigation of the parallel analysis criterion for determining the number of common factors. Multivariate Behav Res 1975; 10(2): 193-206.
16. Montanelli RGJr, Humphreys LG. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: a Monte Carlo study. Psychometrika 1976; 41(3): 341-48.
17. Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometric Monograph 1969:(Suppl. 17):S1-S100
18. Samejima, F. Graded response model. In: van LindenWJ, Hambleton RK, editors. Handbook of modern item response theory. New York: Springer-Verlag; 1997. p. 85-100.
19. Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. Psychometrika 2001; 66(4):507-14
20. Orlando M, Sherbourne, CD, Thissen, D. Summed-score linking using item response theory: application to depression measurement. Psychol Assess 2000; 12(3): 354-9.
21. Nunnally JC, Bernstein IH, Psychometric Theory (3rd Ed.), New York: McGraw Hill; 1994.
22. Carmody TJ, Rush, AJ, Bernstein, IH, Brannan S, Husain MM, Trivedi MH.. Making clinicians lives easier: Guidance on use of the QIDS self-report in place of the MADRS. J Affect Disord 2006; 95(1-3): 115-8.
23. Bernstein IH, Rush AJ, Carmody TJ, Woo A, Trivedi MH. Clinical vs. self-report versions of the quick inventory of depressive symptomatology in a public sector sample. J Psychiatr Res 2007; 41(3-4):239-46.
24. Bernstein IH, Rush AJ, Yonkers K, Carmody TJ, Woo A, McConnell K, et al. Symptom features of postpartum depression: are they distinct? Depress Anxiety 2008; 25(1):20-6.

25. Bernstein IH, Wendt B, Nasr SJ, Rush AJ. Screening for Major Depression in Private Practice. J Psychiatr Pract 2009; 15(2): 87-94.

26. Bernstein IH, Rush AJ, Trivedi MH, Hughes CW, Macleod L, Witte BP, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology in adolescents. Int J Methods Psychiatr Res 2010; 19(4):185-94.

27. Doraiswamy PM, Bernstein IH, Rush AJ, Kyutoku Y, Carmody TJ, Macleod L, et al. Diagnostic utility of the Quick Inventory of Depressive Symptomatology (QIDS-C_{16} and QIDS-SR_{16}) in the elderly. Acta Psychiatr Scand 2010; 122(3):226-34.

28. Bernstein I H, Rush A J, Suppes T, Kyotoku Y, Warden D. The Quick Inventory of Depressive Symptomatology (Clinician and Self-Report Versions) in Patients with Bipolar Disorder. CNS Spectr 2010; 15(6):367-73.

29. Bernstein IH, Rush AJ, Stegman D, Macleod L, Witte B, Trivedi MH. A Comparison of the QIDS-C_{16}, QIDS-SR_{16}, and MADRS in an Adult Outpatient Clinical Sample. CNS Spectr 2010; 15(7): 458-68.

30. Bilgel N, Bayram N. Turkish Version of the Depression Anxiety Stress Scale (DASS-42): Psychometric Properties. [Depresyon Anksiyete Stres Ölçeğinin (DASS-42) Türkçeye Uyarlanmış Şeklinin Psikometrik Özellikleri]. Archives of Neuropsychiatry-Nöropsikiyatri Arşivi Dergisi 2010; 47(2):118-26.

31. Bernstein IH, Rush AJ, Suppes T, Trivedi MH, Woo A, Kyutoku Y, et al. A psychometric evaluation of the clinician-rated Quick Inventory of Depressive Symptomatology (QIDS-C_{16}) in patients with bipolar disorder. Int J Methods Psychiatr Res 2009; 18(2):138-46.

32. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, et al. Screening for depression in adults: A summary of the evidence for the U.S. Preventative Service Task Force. Ann Intern Med 2002; 136(10):765-76.

33. Trivedi MH. Tools and strategies for ongoing assessment of depression: a measurement-based approach to remission. J Clin Psychiatry 2009;70 (Suppl. 6):26-31.

34. Brown ES, Murray M, Carmody TJ, Kennard BD, Hughes CW, Khan DA, et al. The Quick Inventory of Depressive Symptomatology-Self-report: a psychometric evaluation in patients with asthma and major depressive disorder. Ann Allergy Asthma Immunol 2008; 100(5):433-8.

35. Carmody TJ, Rush AJ, Bernstein IH, Brannan S, Husain MM, Trivedi MH. Making clinicians lives easier: guidance on use of the QIDS self-report in place of the MADRS. J Affect Disord 2006; 95(1-3):115-8.

36. Rush AJ, Bernstein IH, Trivedi MH, Carmody TJ, Wisniewski S, Mundt JC, et al. An evaluation of the Quick Inventory of Depressive Symptomatology and the Hamilton Rating Scale for Depression: a Sequenced Treatment Alternatives to Relieve Depression trial report. Biol Psychiatry 2006;59(6): 493-501.

KISA DEPRESİF BELİRTİ ENVANTERİ ÖZBİLDİRİM FORMU

(hasta tarafından doldurulacak)

Son 7 gün boyunca, sizi en iyi tanımlayan seçeneği daire içine alınız.

1. Uykuya dalma:
0 Uykuya dalmam hiçbir zaman 30 dakikayı aşmıyordu.
1 Bu sürenin yarısından azında, uykuya dalmam en az 30 dk. sürüyordu.
2 Bu sürenin yarısından çoğunda, uykuya dalmam en az 30 dk sürüyordu.
3 Bu sürenin yarısından çoğunda, uykuya dalmam 60 dakikadan uzun sürüyordu.

2. Gece boyunca uyku:
0 Gece uyanmıyordum.
1 Her gece kısa sürelerle birkaç kez uyanarak, huzursuz ve hafif uyuyordum.
2 Gecede en az bir kez uyanıyordum, ancak kolayca tekrar uyuyordum.
3 Bu sürenin yarısından çoğunda, gece boyu birden fazla uyanıyordum ve 20 dakika ya da daha uzun süre uyanık kalıyordum.

3. Çok erken uyanma:
0 Bu sürenin çoğunda, kalkmam gereken zamandan en fazla 30 dakika önce uyanıyordum.
1 Bu sürenin yarısından çoğunda, kalkmam gerekenden 30 dk.dan uzun bir süre öncesinde uyanıyordum.
2 Hemen her zaman, gerekenden en az bir saat önce uyanıyordum, ancak sonuçta tekrar uyuyordum.
3 Gerekenden an az bir saat önce uyanıyordum ve bir daha uyuyamıyordum.

4. Çok fazla uyuma:
0 Gün içinde uyuklamaksızın, gecede en fazla 7/8 saat uyuyordum.
1 Gündüz uyuklamalar da dahil olmak üzere 24 saat boyunca, en fazla 10 saat uyuyordum.
2 Gündüz uyuklamalar da dahil olmak üzere 24 saat boyunca, en fazla 12 saat uyuyordum.
3 24 saat boyunca uyuklamalar da dahil olmak üzere, 12 saatten fazla uyuyordum.

5. Keder hissi:
0 Kederli hissetmiyordum.
1 Bu sürenin yarısından azında kederli hissediyordum.
2 Bu sürenin yarısından çoğunda kederli hissediyordum.
3 Bu sürenin hemen hepsinde kederli hissediyordum.

6. İştah azalması:
0 İştahımda her zamankine göre değişiklik olmadı.
1 Her zamankinden daha az miktar ya da sıklıkta yiyordum.
2 Her zamankinden belirgin olarak daha az ve kendimi zorlayarak yiyordum.
3 24 saat içinde nadiren ve yalnızca kendimi çok zorlayarak ya da başkalarının zorlaması ile yiyordum.

7. İştah artması:
0 İştahımda her zamankine göre değişiklik olmadı.
1 Her zamankinden daha sık yeme ihtiyacı duyuyordum.
2 Düzenli olarak, her zamankine göre daha sık ve/veya daha fazla miktarda yiyordum.
3 Hem öğünlerde hem de öğün aralarında aşırı yeme isteği duyuyordum.

8. Kilo verme (son iki hafta içerisinde):
0 Kilomda bir değişiklik olmadı.
1 Hafif bir kilo kaybım olduğunu hissediyorum.
2 1 kilogram ya da daha fazla verdim.
3 2,5 kilogram ya da daha fazla kilo kaybettim.

9. Kilo alma (son iki hafta içinde):
0 Kilomda bir değişiklik olmadı.
1 Hafif kilo aldığımı hissediyorum.
2 1 kilogram ya da daha fazla kilo aldım.
3 2,5 kilogram ya da daha fazla kilo aldım.

10. Konsantrasyon (karar verme):

0 Her zamanki konsantrasyon ve karar verme yeteneğimde bir değişiklik yok.
1 Ara sıra kararsız olduğumu ya da dikkatimin dağıldığını hissediyorum.
2 Çoğunlukla dikkatimi toplamak ya da karar vermek bir çaba gösteriyorum.
3 Okumak için yeterince konsantre olamıyorum ya da basit kararları bile alamıyorum.

11. Kendime bakışım:

0 Kendimi diğerleri kadar değerli ve hak sahibi görüyorum.
1 Her zamankinden daha fazla kendimi suçluyorum.
2 Diğerleri için sorun kaynağı olduğuma büyük ölçüde inanıyorum.
3 Sürekli kendimdeki küçük ya da önemli eksiklikleri düşünüyorum.

12. Ölüm ya da intihar düşünceleri:

0 Ölüm ya da intiharı düşünmüyorum.
1 Hayatın boş olduğunu ya da yaşamaya değip değmeyeceğini düşünüyorum.
2 Haftada birkaç kez birkaç dakika boyunca intihar ya da ölümü düşünüyorum.
3 Günde birkaç kez intihar ya da ölümü bazı ayrıntılarıyla düşünüyorum ya da intihar için özgün planlar yaptım ya da yaşamıma son vermeyi denedim.

13. Genel ilgi:

0 Diğer insanlar ye genel aktivitelere ilgim, her zamankinden farklı değil.
1 Diğer insanlar ye genel aktivitelere ilgimin daha az olduğunu fark ediyorum.
2 Önceki aktivitelerimin yalnızca bir ya da ikisine ilgimin sürdüğünü fark ettim.
3 Önceki aktivitelerime hemen hemen hiç ilgim kalmadı.

14. Enerji düzeyi:

0 Her zamanki enerji düzeyimde bir değişiklik yok.
1 Her zamankinden daha kolay yoruluyorum.
2 Olağan günlük aktivitelerime başlamak ya da bitirmek için büyük çaba göstermem gerekiyor (alışveriş, ev işleri, yemek yapma ve işe gitme gibi).
3 Enerjim olmadığı için olağan günlük aktivitelerimin çoğunu yapamıyorum.

15. Yavaşlama hissi:

0 Her zamanki olağan hızımda düşünüp, konuşup hareket ediyorum.
1 Daha yavaş düşündüğümü ya da sesimin düzeyinin donuk olduğunu fark ediyorum.
2 Soruların çoğuna yanıt vermem birkaç saniye gerektiriyor ve düşüncemin yavaşladığına eminim.
3 Sıklıkla aşırı çaba harcamadan sorulara yanıt veremiyorum.

16. Huzursuzluk hissi:

0 Huzursuz hissetmiyorum.
1 Sık sık huzursuzluk hissediyorum, ellerimi ovuşturuyor ya da oturma biçimimi değiştiriyorum.
2 Hareket etme isteği duyuyorum ye çok huzursuzum.
3 Zaman zaman oturarak bekleyemiyorum ve dolaşma ihtiyacı duyuyorum.

Puanlamak için:

1. Uyku ile ilgili 4 maddeden (1-4) en yüksek puanı seçiniz _____
2. Madde 5 _____
3. İştah ile ilgili 4 maddeden (6-9) en yüksek puanı seçiniz _____
4. Madde 10 _____
5. Madde 11 _____
6. Madde 12 _____
7. Madde 13 _____
8. Madde 14 _____
9. Psikomotor durumla ilgili 2 maddeden (15-16) en yüksek puanı seçiniz _____

Toplam puan: (0-27) _____