

GERİATRİK DEPRESYON ÖLÇEĞİNİN TÜRK YAŞLI NÜFUSUNDA GEÇERLİLİK VE GÜVENİLİRLİĞİ¹

Turan ERTAN², Engin EKER², Vedat ŞAR³

Özet

Gerekeçe: Geriatrik Depresyon Ölçeği'nin 60 yaş üzerindeki Türk popülasyonunda geçerlilik ve güvenilirliğinin incelenmesi bu çalışmanın amacını oluşturmaktadır.

Yöntem: Çalışmada kendi evinde ya da huzurevinde yaşayan 319 denekten oluşan kontrol grubu ile 32 major depresyon hastasından oluşan depresyon grubu olmak üzere iki denek grubu ele alınmıştır.

Sonuçlar: Ölçek toplam puanı yeterli bir test-tekrar test tutarlılığı ($r: 0.77$) göstermiş ve yüksek iç tutarlılık ($\alpha: 0.92$) saptanmıştır. Ölçeğin geçerlilik analizinde ise major depresyon grubu ile kontrol grubu ortalamaları arasında çok ileri düzeyde anlamlı fark ($p < 0.001$) tespit edilmiştir. Ölçeğe ait sorular tek tek ele alındığında, 5-8-11-13-14-16 ve 23'üncü soruların düşük düzeyde test-tekrar test tutarlılığı görmekteyiz. Yine 2-6-20-22-26 ve 30'uncu soruların kontrol grubunda, 12 ve 27'inci soruların major depresyon grubunda ve de 12 sorunun iki denek grubu birarada ele alındığında ölçek toplamı puanı ile düşük düzeyde korelasyon yani düşük iç tutarlılık gösterdiğini saptamaktayız.

Değerlendirme: Bu sonuçlarla ölçeğin Türk yaşlı popülasyonunda depresyon tanısı açısından yeterli geçerlik ve güvenilirliğe sahip olduğu sonucuna varılmıştır. Düşük test-tekrar test tutarlılığı veya düşük soru-toplam skor korelasyonu gösteren soruların ölçekten çıkarılması ile geçerlilik ve güvenilirliği daha yüksek bir kısa formun elde edilebileceğini düşündürmektedir.

Anahtar kelimeler: Depresyon, Ölçek, Geriatrik depresyon ölçeği, Geçerlilik, Güvenilirlik

RELIABILITY AND VALIDITY OF THE GERIATRIC DEPRESSION SCALE IN TURKISH ELDERLY POPULATION

Summary

Objective: Reliability and validity of the Geriatric Depression Scale was investigated in Turkish population over 60 years of age.

Methods: 32 patients with major depression and 319 pensioners living at home or in a state residency participated in the study.

Results: A high correlation level ($r: 0.77$) was obtained for the scale total score. The scale also showed high internal consistency ($\alpha: 0.92$). The validity analysis of the scale resulted in a very high significant difference ($p < 0.001$) between total scores of the control and depressed groups. Whereas seven items (5-8-11-13-14-16-23) showed low correlation ($r < 30$) in the test-retest analysis. On the other hand, six items (2-6-20-22-26-30) in the control group and two items (12-27) in the depressed group had low item-total score correlation ($r < 30$). One item (12) had low item-total score correlation while the analysis was repeated for both groups together.

Discussion: The results of the study indicated that the scale has reliability and validity for the Turkish elderly population and a short version with higher reliability and validity could be obtained by eliminating the item that had low test-retest or item-total score correlation score.

Key words: Depression, Screening test, Geriatric depression scale, Elderly, Reliability, Validity

Giriş

Depresyon geriatrik psikiyatri alanında en sık karşılaşılan sorunlardan biridir. Batı toplumlarında yapılan depresyon prevalans çalışmaları yaşlı grupta depresyon prevalansı % 11.3 olarak bildirilmekte ve (1) major depresyon % 4.6, distimik bozukluk % 6.7 sıklığında bulunmaktadır (2). Evinde

de yaşayan 70 yaş üzerindeki nüfusta major depresyon prevalansı % 6.3 iken bu oran 80 yaş üzeri grupta % 15.5'e yükselmektedir (3). Bugün kadar Türkiye'de yaşlı nüfusta depresyon prevalansı üzerine yayımlanmış örneğine ulaşabildiğimiz tek çalışma Bekaroğlu ve ark. tarafından Trabzon bölgesinde gerçekleştirilmiştir (4). Yazarlar kendi

1 32. Ulusal Psikiyatri Kongresi'nde (25-28 Eylül 1996, Ankara) sözel bildiri olarak sunulmuştur.

2 İ.Ü. Cerrahpaşa Tıp Fakültesi Psikiyatri Anabilim Dalı, Geropsikiyatri Bilim Dalı

3 İ.Ü. İstanbul Tıp Fakültesi Psikiyatri Anabilim Dalı

evinde yaşayan 60 yaş üzeri nüfus içinde, major depresyon tanısı alan hastaların sıklığını % 29 olarak ileri sürmektedirler. Görüldüğü gibi yaşlılık dönemi depresyonları gerek batı toplumlarında gerekse Türkiye'de sık karşılaşılan bir tablodur.

Yaşlılık depresyonunun değerlendirilmesi diğer yaş gruplarında görülen depresif atakların değerlendirilmesinden farklı bir yaklaşım gerektirmektedir. Somatik yakınmalar yaşlılarda gençlere göre daha sık olarak depresyon dışı nedenlere bağlı olabilmekte ve bu yaş grubunda depresyon tanısı açısından gençlerde olduğundan daha az değer taşımaktadır. Örneğin uyku bozuklukları depresyonda sıklıkla ortaya çıkan bir belirti olmakla birlikte depresif olmayan yaşlılarda da sıkça karşılaşılan bir durumdur (5). Cinsel işlevde azalma, kabızlık ve sıklıkla eklem patolojilerine bağlı olabilen ağrılar psikiyatrik açıdan sağlıklı yaşlılarda nadir olmayarak görülebilmektedir. Somatik yakınmaların yaşlılarda depresyon dışı nedenlere bağlı olmasının yanı sıra, yaşlıda var olan bedensel hastalığa bağlı bedensel yakınmaların hekim açısından depresyon tanısı koymakta güçlük yaratması ve sonuçta tanı alması gereken depresyonlu hastaların hekimlerin gözünden kaçması da mümkündür. Fiziksel hastalık nedeniyle hastanelerde yatan yaşlı hastalarda tanısı konmamış depresyon vakalarının tanımlandığı çalışmalara rastlamak mümkündür (6). Bu nedenlere bağlı olarak genç erişkin bireylerde depresyon tanısı koymak için kullandığımız ölçütler yaşlıda yeterli ve/veya güvenilir olmayabilir. Yaşlı nüfus için, içerdikleri soruların bedensel belirtilere daha az yer verdiği ve mümkünse öz bildirime dayalı, yaşlı için kolay anlaşılabilir ve yanıtlanır sorulardan oluşan, genel nüfus için kullanılanlardan daha farklı depresyon ölçeklerinin kullanılması gereklidir.

Türkiye'de yaşlılık depresyonunun değerlendirilmesinde objektif araç olarak kullanılabilecek herhangi bir ölçek bulunmamaktadır. Batı toplumlarında

kullanılan bu tür ölçeklerin Türk toplumunda sadece çevirisi yapılarak kullanılması mümkündür. Ancak Türk toplumu ekonomik, dinsel ve kültürel alanlarda batı toplumlarından birçok farklılıklar gösterdiğinden, böyle bir ölçeğin Türkiye'de yaşayan yaşlılarda depresyonun değerlendirilmesinde ne derece özgün ve duyarlı olacağından emin olunması bu alanda yapılacak araştırmaların sağlıklılığı açısından kaçınılmazdır.

Bu çalışmanın amacı, Yesavage, Brink ve arkadaşları tarafından hazırlanıp geriatrik nüfus için özgün kabul edilip birçok ülkede çevirisi yapılarak kullanılan Geriatrik Depresyon Ölçeği'nin (GDÖ) (8) Türk yaşlı nüfusunda depresyon tanısı koymaktaki geçerlilik ve güvenilirliğinin incelenmesidir.

ÖLÇEKLE İLGİLİ BİLGİLER

GDÖ Yesavage, Brink ve arkadaşları tarafından oluşturulup 1983 yılında geçerlilik ve güvenilirlik incelemesi yapılmış olan yaşlı nüfusa yönelik bir depresyon ölçeğidir (8). Öz bildirime dayalı 30 sorudan oluşan bu ölçeğin hazırlanışındaki temel amaç yanıtlanması kolay ve yaşlılara yönelik sorulardan oluşmasıdır. Bu nedenle yazarlar, başka uyku bozuklukları, cinsel işlev bozukluğu, vücutta ağrı ve sızılar gibi somatik belirtiler olmak üzere yaşlı nüfus için depresyon dışı nedenlerle de oluşabilecek belirtilerin yer almadığı, yanıtları sadece evet ya da hayır şeklinde olan bir ölçek oluşturmuşlardır (8). Ölçeğin puanlamasında her soruya depresyon lehine verilen yanıt için 1 puan diğer yanıt için 0 puan verilmekte ve sonuçta toplam puan depresyon puanı olarak kabul edilmektedir. Depresyon tanısı için eşik değer (cutt-off point) 11 olarak alındığında ölçeğin duyarlılığı % 85, özgüllüğü ise % 95 olarak hesaplanmıştır (7). Yazarlar eşik eğer 14 alındığında ise duyarlılığın % 80'e düştüğünü, özgüllüğün ise, tüm sağlıklıların depresiflerden ayrılması anlamına gelen % 100'e yükseldiğini saptamışlar ve eşik değerleri, pua-

nın 0-11 arasında olması nedeniyle depresyon olmadığı, 11-14 arası için olası depresyon, 14 ve yukarısı için ise kesin depresyon tanısı yönünde açıklamışlardır. GDÖ'nün oluşumunda ilk aşama kullanılan 100 sorunun yöneldiği alanları somatik belirtileri duygulanım değişimi, bilişsel işlev bozukluğu, güdülenme düşüklüğü, gelecek yönelimi kaybı ve kendilik değeri kaybı olarak sıralanabilir (8). Ölçeğin 30 soruluk son şekli ise duygulanım değişimi, durağanlaşma, çabuk öfkelenme, yaşamdan geri çekilme (withdrawal), rahatsız edici düşünceler ve geçmiş, şimdiki zaman ve geleceğe yönelik olumsuz yargı alanlarını içermektedir. Ölçeğin geçerliliği ve güvenilirliği toplumda yaşayanlarda (8), psikiyatri kliniklerinde yatan hastalarda (9), bakım evlerinde yaşayanlarda (10) ve ayakta tedavi gören hastalarda (11) değerlendirilmiş ve geçerli olduğu saptanmıştır. Bunun yanısıra tıbbi hastalık nedeniyle tedavi altında olanlarda (6), ve demanslılarda da (12) ölçek geçerli bulunmuştur. Koenig ve arkadaşları tıbbi hastalığı olan depresyonlu denekleri, aynı grubun depresyonu olmayan deneklerinden ayırması açısından GDÖ'nün 14 puanlık eşik değere göre, % 86 özgüllük, % 93 duyarlılık, % 63 pozitif öngörü değeri (ölçek tarafından tespit edilen depresyonlu deneklerin varolan depresyonlu denek sayısına oranı), % 98 negatif öngörü değeri (ölçek tarafından depresif olmayan olarak tanımlanan denek sayısının varolan tüm depresif olmayan denek sayısına oranı) ile güvenilir sonuç verdiği-

ni saptamışlardır (6). Lichtenberg ve arkadaşları ise ölçeği demans hastalarında depresyon tanısı koymadaki duyarlılığı açısından Hamilton depresyon ölçeği ile karşılaştırmışlar ve GDÖ'yü % 82 duyarlılık ve % 88 özgüllük ile Hamilton'a göre üstün bulmuşlardır (12).

DENEK GRUPLARI VE YÖNTEM

DENEK GRUPLARI

Bu çalışmanın deneklere uygulanması 1992-1995 yılları arasında yapılmıştır. Çalışmaya denek grubu olarak yaşları 60 ile 93 arasında değişen, 216'sı kendi evinde yaşayan, 103'ü T.C. Emekli Sandığı İstanbul Etiler Huzurevi'nde kalan toplam 319 kişiden oluşan kontrol grubu ve İ.Ü. Cerrahpaşa Tıp Fakültesi Geropsikiyatri Bilim Dalı yataklı servisinde ya da polikliniğinde DSM-IV tanı ölçütlerine göre major depresyon tanısı konulmuş ve genel sağlığı ya da yaşamını etkileyebilecek herhangi bir organik hastalığı bulunmayan 32 hastadan oluşan major depresyon grubu olmak üzere iki grup alındı. Kendi evinde yaşayan denekler emekli maaşlarını almak üzere İstanbul ili çeşitli semtlerindeki bankalara müracaat eden kişiler arasından rastgele seçildi. Tüm deneklerde en az ilkokul mezunu olma koşulu arandı. Denek seçiminde herhangi bir psikiyatrik test uygulanmaksızın, klinik değerlendirmeye dayanılarak bilişsel işlev bozukluğu saptananlar çalışmaya alınmadı. Kendi evinde ve huzurevinde

Tablo 1. Denek grupları cinsiyet ve yaş dağılımları

	Majör depresyon grubu n: 32	Kontrol grubu n: 319	
KADIN	%62.5 (n:20)	%44.2 (n:141)	X ² =7.34, df 1, p 0.11
ERKEK	%37.5 (n:12)	%55.8 (n:178)	
60-69 YAŞ	%65.5	%56.1	X ² =1.06, df 2, p 0.58
70-79 YAŞ	%27.6	%31.2	
≥ 80 YAŞ	%10.3	%12.7	
Ortalama yaş	68.34 ± 6.49	69.50 ± 9.31	t=0.6 df 341, p 0.51
Minimum yaş	60.00	60.00	
Maksimum yaş	83.00	93.00	

yaşayanlarda psikopatoloji açısından başka bir değerlendirme yapılmadı. Denek gruplarının cinsiyet ve yaş dağılımları Tablo 1'de görülmektedir.

Major depresyon grubu hastalarının İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Geropsikiyatri Bilim Dalı uzman ve asistanları tarafından değerlendirmelerinde DSM-IV tanı ölçütleri dışında herhangi bir depresyon değerlendirme ölçeği kullanılmadı. Ölçeğin doldurulmasında deneklerin kendilerinin okuyarak doldurması esasına göre uygulama yapılırken sadece deneklerin gözlüğünün yanında olmaması durumlarında araştırmacılar tarafından sorular hiçbir yorum ya da açıklama yapılmaksızın okunarak yardımcı olundu. Denek tarafından anlaşılmadığı ifade edilen sorular ikinci kez okundu ve tekrar anlaşılamaması halinde boş bırakıldı. Ölçeklerin denekler tarafından doldurulması gerektiğinden tüm deneklerin seçiminde okur yazar olma koşulu arandı. Ölçeklerdeki sorulardan herhangi birini boş bırakılmış olması durumunda o ölçek değerlendirilmeye alınmadı.

YÖNTEM

Ölçeğin Türkçe'ye çevirilmesi

Ölçeğin İngilizce orijinalinden Türkçe'ye çevirilmesi her iki dili de akıcı olarak konuşan iki ayrı psikiyatrist tarafından yapıldı. Bu aşamada psikiyatristlerden birincisi ölçeği İngilizceden Türkçeye çevirdikten sonra oluşan Türkçe çeviriyi ikinci psikiyatrist tekrar İngilizceye çevirdi ve orijinal İngilizce versiyon ile çeviri İngilizce versiyon karşılaştırıldı. Her iki versiyonun birbiri ile uyumlu olduğu sonucuna varıldı.

Ölçeğin dilinin anlaşılabilirliğinin denemesi amacıyla, ön çalışma olarak 30 kişiye uygulanması ve sonuçlara göre değişikliklerin yapılması

İlk olarak 30 kişilik rastgele seçilmiş bir gruba uygulanan pilot çalışmada deneklerin anlamakta zorluk çektikleri sorular tekrar değerlendirilerek, orijinal anlamına bağlı kalınmak koşuluyla bazılarında ifade değişiklikleri yapıldı. Orijinal şekilleriyle anlamı

Tablo 2. GDO Test-tekrar test tutarlılık incelemesi (Pearson ve Spearman korelasyon analizi sonuçları, Pearson*) (n: 31)

Soru No	Korelasyon değeri	Soru no.	Korelasyon değeri
Toplam puan	r:0.77* p:0.000		
Soru 1	r:0.47 p:0.008	Soru 16	
Soru 2	r:0.30 p:0.105	Soru 17	r:0.46 p:0.010
Soru 3	r:0.40 p:0.028	Soru 18	r:0.34 p:0.061
Soru 4	r:0.39 p:0.030	Soru 19	r:0.67 p:0.000
Soru 5	r:0.05 p:0.755	Soru 20	r:0.52 p:0.003
Soru 6	r:0.59 p:0.001	Soru 21	r:0.72 p:0.000
Soru 7	r:0.53 p:0.002	Soru 22	r:0.48 p:0.007
Soru 8	r:-0.08 p:0.640	Soru 23	r:0.11 p:0.559
Soru 9	r:0.52 p:0.003	Soru 24	r:0.65 p:0.000
Soru 10	r:0.87 p:0.000	Soru 25	r:0.64 p:0.000
Soru 11	r:0.14 p:0.432	Soru 26	r:0.32 p:0.078
Soru 12	r:0.47 p:0.008	Soru 27	r:0.55 p:0.001
Soru 13	r:0.25 p:0.167	Soru 28	r:0.65 p:0.000
Soru 14	r:0.42 p:0.020	Soru 29	r:0.52 p:0.003
Soru 15	r:0.42 p:0.020	Soru 30	r:0.86 p:0.000

Türkçe'de farklı algılandığı saptanan "gelecekte umutlu musunuz?" şeklindeki soru (soru 5, "Are you hopeful about the future?") olumludan "Gelecekte umutsuz musunuz?" şeklinde olumsuz şekile çevirildi. Bunun karşılığında ölçekteki olumsuz ve olumlu soru sayısının orijinal ölçeği uyumlu olması amacıyla, olumlu ya da olumsuz şeklinin anlam farklılığı göstermediği saptanan bir soru, aynı şekilde ters anlama çevirildi. Bu amaçla 2'inci soru olan "Have you dropped many of your activities and interests?" Türkçe'de ters anlamına gelen "Kişisel etkinlik ve ilgi alanlarımızın çoğunu halen sürdürüyor musunuz?" şeklinde kullanıldı. Söz konusu pilot çalışma sonunda 30 soruluk GDÖ İngilizce formunun Türkçe karşılığı oluşturuldu (Tablo 2).

Test-tekrar test tutarlılığı

Ölçeğin test-tekrar test tutarlılık çalışması Etiler Huzurevinde yaşayan ve huzurevi psikologları tarafından gözlem altında olup bilişsel işlev bozukluğu olmayan 31 kişilik bir grupta gerçekleştirildi. Aynı deneklere ölçek bir kez uygulandıktan sonra 7 gün sonra tekrar uygulandı. Toplam puan ve her bir sorunun 7 gün arayla tutarlılığı Pearson ve Spearman Korelasyon hesaplamaları ile değerlendirildi.

İç tutarlılık

Tüm ölçek ve her bir sorunun ölçek içi tutarlılığı tüm denek grubu, kontrol grubu ve depresyon grubunda Cronbach's alpha ve Split Half testleri ile incelendi.

Geçerlilik

Kolmogorov-Smirnov testine göre çalışmada kullanılan iki grubun ölçek toplam puanlarının normal dağılım göstermemeleri nedeniyle ölçeğin gruplar arasında ayıricılığının ölçülmesi nonparametrik yöntem olarak Mann-Witney-U testi ile yapıldı.

Bulgular

Tekrar-test çalışmasında yer alan huzurevinde yaşayan 31 kişilik gru-

bun yaş dağılımı 65 ve 87 arasında değişmekte olup ortalama yaş 75.53 ± 5.71 'dir. Bu grubun ortalama eğitim süresi ise 11.03 yıldır. Bu grubun % 90.3'si kadınlardan % 9.7 ü erkeklerden oluşmaktadır.

Test-tekrar test duyarlılığı

Ölçeğin birer hafta arayla, huzurevinde yaşayan 31 kişilik denek grubuna uygulanması sonrasında, toplam puan ve her bir sorunun deneklerin bir hafta sonra verdiği yanıtlarla Pearson ve Spearman Korelasyon yöntemi ile karşılaştırmasından elde edilen sonuçlar Tablo 3'deki gibidir. Bu sonuçlara göre GDÖ Türkçe versiyonu toplam puanının aynı kişilere 7 gün arayla yapılan uygulama sonucunda 0.77 düzeyinde orta derecede korelasyon göstererek tutarlı olduğunu görmekteyiz. Tek tek soruların ikinci uygulama ile korelasyonlarına bakıldığında ise 5, 8, 11, 13, 14, 16 ve 23 no'lu soruların 0.30 değerinin altında bir değerle düşük düzeyde korelasyon gösterdiği görülmektedir.

İç tutarlılık

Ölçeğin iç tutarlılık incelenmesi Cronbach's alfa ve Split-Half yöntemlerine göre yapılmıştır. Sonuçlar Tablo 4'de görülmektedir.

Ölçek genel alfa ve split half değerleri ölçüt olarak alındığında GDÖ Türkçe versiyonu gerek kontrol, gerek depresyon grubunda, gerekse her iki grup bir arada ele alındığında, hem alfa değeri hem de Split-Half korelasyon değerine göre 0.90 dolayında çok yüksek bir iç tutarlılık göstermektedir. Ancak tek tek her bir sorunun ölçek toplam puanı ile korelasyonuna bakıldığında, 2, 6, 20, 22, 26 ve 30 no'lu soruların depresyon grubunda, 12 ve 27 no'lu soruların kontrol grubunda ve yine 12 no'lu sorunun iki grup bir arada ele alındığında düşük değerlerde korelasyon gösterdiğini görmekteyiz.

Geçerlilik

Ölçek toplam puan ve her bir sorunun depresyon grubunu kontrol grubundan ayırıp ayıramadığına dair

Tablo 3. GDÖ iç-tutarlılık incelemesi

	T0m gruplar sh: 9210	n: 351 alpha: 9270	Kontrol sh: 9036	n: 319 alpha: 9139	Depresyon sh: 9197	n: 32 alpha: 8919
Soru	soru toplam puan korelasyon değeri	Ölçeğin soru harici alpha değeri	Soru toplam puan korelasyon değeri	Ölçeğin soru harici alpha değeri	Soru toplam puan korelasyon değeri	Ölçeğin soru harici alpha değeri
1	.5139	.9248	.4720	.9115	.5667	.8857
2	.4147	.9261	.3773	.9130	.0986	.8952
3	.6810	.9225	.6650	.9083	.5087	.8874
4	.6440	.9230	.6490	.9085	.4345	.8888
5	.4099	.9262	.3752	.9131	.5745	.8855
6	.5154	.9247	.5068	.9109	.1411	.8948
7	.6141	.9234	.5668	.9100	.6669	.8840
8	.5217	.9247	.4645	.9116	.4137	.8891
9	.5898	.9238	.5267	.9107	.6975	.8829
10	.7478	.9216	.7220	.9075	.5498	.8870
11	.6695	.9227	.6385	.9088	.5087	.8874
12	.2889	.9278	.2202	.9056	.6336	.8943
13	.6253	.9233	.5776	.9098	.6944	.8834
14	.4723	.9253	.4542	.9117	.3942	.8897
15	.4653	.9255	.3702	.9130	.3942	.8897
16	.6323	.9232	.6002	.9094	.6395	.8845
17	.6525	.9230	.5998	.9096	.5241	.8870
18	.5481	.9243	.5630	.9100	.4547	.8883
19	.5690	.9241	.5347	.9105	.4083	.8892
20	.3744	.9266	.3743	.9131	.0827	.8951
21	.5327	.9245	.4922	.9112	.4019	.8893
22	.6664	.9227	.6775	.9082	.2005	.8934
23	.3709	.9267	.3219	.9140	.3279	.8906
24	.5594	.9242	.5459	.9103	.4019	.8893
25	.5550	.9242	.5390	.9104	.5275	.8846
26	.5752	.9247	.5063	.9109	.1944	.8939
27	.4463	.9256	.2947	.9136	.5032	.8872
28	.4034	.9262	.3086	.9140	.7491	.8818
29	.4198	.9860	.3850	.9129	.5145	.8869
30	.3670	.9268	.3477	.9136	.0426	.8958

olarak anlamlı derecede ayırmaktadır. Bu şekliyle ölçeğin Türk yaşlı nüfusunda geçerli olduğunu ileri sürebiliriz.

Tartışma

30 soruluk GDÖ test-tekrar test tutarlılığı incelemesinde, 7 gün arayla aynı kişilere yapılan iki uygulamada iki test toplam puanları arasında 0.77 düzeyinde pozitif korelasyon bulunmaktadır (Tablo 3). Bu sonuç ölçeğin bütünü için yeterli kabul edilebilir ve Türk olmayan toplumlarla yapılan incelemelerde alınan sonuçlara benzerlik göstermektedir. Yesavage ve arkadaşları GDÖ'nün orijinal şeklini tanıttıkları makalelerinde (8) 20 denek kullanı-

puanının iki uygulama arasındaki korelasyonunu 0.98 olarak saptamışlardır. Görüldüğü gibi özgül olmayan bir denek grubu kullanmış olan Yesavage ve ark.'nın sonuçları bizim sonucumuza daha yakındır. Lyons ve ark.'nın çalışmasında ise kalça kırığı hastalarından oluşan özgül bir grubun kullanılmış olması bu çalışmanın Yesavage ve ark. ile bizim çalışmamızdan farkını oluşturmaktadır. Deneklerin sabit bir fiziksel engelinin bulunmasının ölçek üzerinde bir belirleyiciliği olmuş olabilir. Ayrıca Lyons ve arkadaşları hastaların fiziksel sorunu bulunan ameliyat edilmiş hastalar olmaları nedeniyle ölçeği 2-3 aşamada dinlenerek ya-

Tablo 4. GDO toplam puan ve herbir sorusunun geçirgenliğinin Mann-Witney U-Wilcoxon testi ile incelenmesi (U: U değeri, W: Wilcoxon değeri)

	Kontrol grubu ort. sıra değeri (n: 319)	Depresyon grubu ort. sıra değeri (n: 32)	U	W	Z	P
1	148.82	196.60	2847.0	5898.0	-3.7377	(.002)
2	147.49	208.80	2481.0	6264.0	-4.2390	(.000)
3	147.14	212.00	2385.0	6360.0	-4.4243	(.000)
4	148.92	195.60	2877.0	5868.0	-3.2231	(.013)
5	150.03	185.40	3183.0	5562.0	-2.4286	(.052)
6	148.40	200.40	2733.0	6012.0	-3.7262	(.002)
7	147.00	213.30	2346.0	6399.0	-4.6465	(.000)
8	146.58	217.20	2229.0	6516.0	-5.1519	(.000)
9	146.86	214.60	2307.0	6438.0	-5.0145	(.000)
10	145.34	228.60	1887.0	6858.0	-5.8353	(.000)
11	146.27	220.00	2145.0	6600.0	-5.1112	(.000)
12	149.68	188.60	3087.0	5658.0	-2.6464	(.008)
13	146.95	213.80	2331.0	6414.0	-4.6935	(.000)
14	149.52	190.10	3042.0	5703.0	-3.0229	(.025)
15	146.70	216.10	2262.0	6483.0	-7.1433	(.000)
16	147.54	208.30	2496.0	6249.0	-4.1954	(.000)
17	145.20	229.90	1848.0	6897.0	-6.2702	(.000)
18	150.57	180.50	3330.0	5415.0	-2.0326	(.042)
19	147.24	211.10	2412.0	6333.0	-4.6335	(.000)
20	150.20	183.90	3228.0	5517.0	-2.3365	(.019)
21	147.13	212.10	2382.0	6363.0	-4.4129	(.000)
22	147.39	209.70	2454.0	6291.0	-4.3905	(.000)
23	148.40	200.40	2733.0	6012.0	-3.5302	(.000)
24	148.65	198.10	2802.0	5943.0	-3.3954	(.001)
25	149.10	194.00	2925.0	5820.0	-3.1086	(.019)
26	148.28	201.50	2700.0	6045.0	-3.7799	(.002)
27	144.80	233.50	1740.0	7005.0	-8.5802	(.000)
28	147.67	207.10	2532.0	6213.0	-4.2299	(.000)
29	149.65	188.90	3078.0	5667.0	-2.7171	(.006)
30	148.95	195.40	2883.0	5862.0	-3.1579	(.016)
Toplam puan	142.62	253.60	1137.0	7608.0	-6.5308	(.000)

nıtlamalarını sağlamışlardır. 30 sorudan oluşan ve bizim çalışmamız sırasında birçok denek tarafından uzun bulunan GDÖ'nün dinlenerek yanıtlanması bu çalışmadaki yüksek korelasyona neden olmuş olabilir. Ancak sonuçlardan görüldüğü gibi, diğer çalışmalarda olduğu gibi bizim çalışmamızda da korelasyon düzeyi yeterlidir.

Bizim çalışmamızın önemli bir farkı ölçeğin gerek test-tekrar test, gerek iç tutarlılık ve gerekse geçerlilik incelemelerinde sadece toplam puanlara değil her bir soruya ait sonuçlara da dikkat edilmiş olmasıdır. 30 soruluk GDÖ'nün Türkçe versiyonunda 7 sorunun (5, 8, 11, 13, 14, 16 ve 23) iki uygulama arasında düşük korelasyon gösterdiği gözlenmektedir. Literatürde

test-tekrar test incelemesinin her bir soru için bakıldığı nadir çalışmalardan biri olan Burke ve arkadaşlarının çalışmasında (14), GDÖ'nün telefonla yaklaşık 12 gün arayla 101 deneye yapılan iki uygulaması arasında tek tek soruların uyumuna kappa yöntemi ile bakılmış ve 0.26-0.64 arasında değişen değerler elde edilmiştir. Bu çalışmada 3 soruya ait değerler (soru 8, 13 ve 29) kappa yöntemi için orta derecede uyum kriteri olarak kabul edilen 0.4 değerinin altında kalmaktadır. Bu sorulardan iki tanesi (soru 8 ve 13) bizim çalışmamızda da düşük korelasyon göstermişlerdir. İki çalışma arasındaki temel fark Burke ve arkadaşlarının denek gruplarının bizimkine göre daha geniş olmasıdır (31 ve 101).

Sonuç farklılığının bu nedenden kaynaklanıyor olması düşünülebilir ancak korelasyon tipi istatistiksel incelemelerde daha geniş denek grubunun kullanılması tercih edilmekle birlikte 30 denek genellikle yeterli kabul edilmektedir. Ayrıca bu noktada Amerikan ve Türk toplumları arasındaki kültürel farklılıkların bu tür sonuçlara yol açabileceği de düşünülebilir. Kaldı ki farklı kültürlerde farklı sonuçlar elde edilmesi nedeniyle psikiyatrik değerlendirme ölçeklerinin kullanıldıkları toplumlarda geçerlilik ve güvenilirliğinin sınanması bu açıdan gerekli görülmektedir. Bu açıdan sonuçları etkileyebilecek bir diğer faktör de ölçeğin evet/hayır şeklinde yanıtlardan oluşması olabilir. Uygulama sırasında deneklerin bazı sorulara ne evet ne de hayır yanıtını uygun gördüklerini, iki yanıtın birini tercih etmediklerini gözlemledik. Bu denekler araştırma ekibinin uyarısı ile yanıtlardan birini seçmiş bulunmaktadır. Bu durumda bazı deneklerin ikinci uygulamada kendi durumları değişmemesine rağmen diğer yanıtı tercih etmiş olmaları olasılık dahilindedir ve bu durum da sonucu etkilemiş olabilir.

Bizim çalışmamızda, her iki grup bir arada ele alındığında, GDÖ'nün orijinal şekli ile yapılan analizde tüm test alpha değeri 0.92, soru toplam puan korelasyon değerleri ortalama 0.52 olup 0.29-0.75 arasında değişmektedir (Tablo 4). Yesavage ve arkadaşlarının (8) bizim denek grubumuza göre daha dar bir grupta yaptıkları incelemede (normal ve depresif kişilerin birarada ele alındığı 100 kişilik grupta) tüm test alpha değeri 0.94 olarak saptanmıştır. Yazarlar denek gruplarında (açıklıkla belirtilmemiş olmakla birlikte makaleden depresif ve depresif olmayan grupların birlikte ele alındığı anlaşılmaktadır) soru-toplam puan korelasyonlarının 0.32-0.83 arasında değiştiğini, ortalamasının ise 0.56 olduğunu belirtmektedirler. Dunn ve Sacco'nun çalışmalarında ise (15) toplum içinde yaşayan 196 denekle yapılan uygulamada testin genel alpha değeri

0.91, median soru-toplam puan korelasyonu 0.48 bulunmuş, değerlerin alt ve üst sınırları 0.25 ile 0.69 arasında tespit edilmiştir. Ancak soruların tek tek toplam puan ile korelasyon değerleri belirtilmemiştir. Bu alanda yapılmış olan bir diğer çalışmada ise Cialdella ve arkadaşları Fransız toplumunda 55'i depresif olmayan 44'ü depresif olan toplam 99 kadın denekle yaptıkları incelemede (16) tüm test alpha değerini 0.90, ortalama soru-toplam puan korelasyonunu ise 0.47 olarak saptamıştır. Bu sonuçlar değişik çalışmalar ile bizim çalışmamız arasında uyumluluk bulunduğunu göstermektedir. Bizim denek grubumuzun Yesavage ve arkadaşlarının grubundan farklılığı, yazarların 100 kişilik grubunun 40 depresif olmayan ve 60 depresif denekten, bizimkinin ise 319 kontrol denegi ve 32 depresif hastadan oluşmasıdır. İki çalışma arasındaki temel fark, bizim kontrol grubumuzun sadece bilişsel işlev yeterliliği aranmış, muhtemelen içerisinde depresyonluların da bulunduğu bir grup olmasıdır. GDÖ'nün Türkçe versiyonunun Yesavage ve arkadaşlarının geçerlilik analizi sonuçları ile karşılaştığımızda, ölçeğin toplam puanının bizim çalışmamızda da yazarların çalışmasında olduğu gibi (14) depresyonlu grubu kontrol grubundan çok ileri derecede anlamlı olarak ayırdığını görmekteyiz (Tablo 5). Yesavage ve arkadaşları normal grup olarak psikiyatrik değerlendirmeden geçirilmiş depresif olmayan kişiler seçmişlerdir. Bu fark bizim çalışmamızın sağlamlığı açısından bir dezavantaj oluşturmamaktadır. Zira kontrol grubu içerisinde muhtemelen bulunan depresyonlu deneklerin almış oldukları puanların etkisine rağmen iki grup arasında çok ileri derecede anlamlı fark saptanmıştır.

Bu çalışmada kontrol grubunun psikiyatrik değerlendirmeden geçirilmiş sağlıklı deneklerden değil toplumda yaşayan yaşlı nüfustan oluşmuş olması nedeniyle eşik değer hesaplamaları yapılmamıştır. Çalışmanın bu aşaması halen kliniğimizde devam etmektedir.

Sonuç

Yaptığımız geçerlilik ve güvenilirlik analizi sonuçlarına göre, GDÖ'nün 60 yaş üzerindeki Türk nüfusunda yeterli test-tekrar test ve iç-tutarlılık güvenilirliğine ve de geçerliliğine sahip olduğunu görmekteyiz.

Ölçeğe ait sorular tek tek ele alındığında, 5-8-11-13-14-16 ve 23'üncü sorularda düşük düzeyde test-tekrar test tutarlılığı görmekteyiz. Yine 2-6-20-22-26 ve 30'uncu soruların kontrol grubunda, 12 ve 27'inci soruların major depresyon grubunda ve de 12. sorunun iki denek grubu birarada ele alındığında ölçek toplam puanı ile düşük düzeyde korelasyon yani düşük iç tutarlılık gösterdiğini saptamaktayız. Bu sorular gerek hastaların tedavi takibinde ve özellikle depresyonu çok şiddetli olmayan hastaların saptanmasında yanıltıcı sonuçlara yol açabilecektir. Bu soruların ölçekten çıkarılması ile oluşabilecek kısa versiyonun ise, Türk toplumunda geçerlilik ve güvenilirliği daha yüksek, kısa olmasının da katkısıyla daha rahat uygulanır olabileceğini düşünebiliriz.

KAYNAKLAR

1. Copeland JRM, Dewey ME, Wood N., Searle R., Davidson LA, McWilliams C, Range of mental illness among the elderly in the community: prevalence in Liverpool using the GMS-AGECAT package. Br. J. Psychiatry 1987, 150, 815-823.
2. Copeland JRM, Gurland BJ., Dewey ME, Keller MJ, Smith AMR, Davidson LA. Is there more dementia, depression and neurosis in New York? Br. J Psychiatry 1987. 151, 466-473.
3. Kay, D.W.K., Henderson A.S., Scott, R., Wilson, J., Rickwood, D. and Grayson, D.A. The prevalence of dementia and depression among the elderly living in the Hobart community: the effect of the diagnostic criteria on the prevalence rates. Psychol. Med. 1985, 15. 771-788.
4. Bekaroğlu M., Uluutku N., ve ark., Depressi-

- on in elderly population in Turkey. Acta Psychiatr. Scand 1991; 84, 174-178.
5. Coleman, R.M., Miles, L.E., Gulleminault, C., Zarcone, V.P., van der Hoed J. Dement, W.C., Sleep-wake disorders in the elderly: a polysomnographic analysis. J. Am. Geriatr. Soc. 1981, 29, 289-296.
6. Koenig HJ, Meador KG., Cohen HJ., Blazer DG., Self rated depression scale and screening major depression in the older hospitalized patient with medical illness. J. Am. Geriatr. Soc 1988, 36, 699-706.
7. Brink T.L., Yesavage J.A., Owen L. ve ark., Screening tests for geriatric depression. Clin Gerontol 1982, 1: 37-43.
8. Yesavage J.A., Brink T.L., Rose T.L. ve ark. Development and validation of a geriatric depression screening scale: a preliminary report. J. Psychiatric Research 1983, Vol. 17, No. 1. pp. 37-49.
9. Hyer L. ve Blount J., Concurrent and discriminant validity of the Geriatric Depression Scale with older psychiatric inpatients. Psychological Reports, 1984. 54, 611-616.
10. Leshner E.L., Validation of the Geriatric Depression Scale among nursing home residents Clinical Gerontologist, 1986, 4, 21-28.
11. Zgoruides G., Spofford M., Doppelt L., The Geriatric Depression Scale, discriminant validity and elderly day-treatment clients. Psychiatric Reports 1989. 64, 1082.
12. Lichtenberg A., Marcopulos B.A., Steiner D.A., Tabscott J.A., Comparison of the Hamilton Depression Rating Scale, and the Geriatric Depression Scale: detection of depression in dementia patients, Psychological reports. 1992, 70, 515-521.
13. Lyons J.S., Strain J.J., Hammer J.S., Ackerman A.D., George F., Reliability, validity and temporal stability of the Geriatric Depression Scale in hospitalized elderly, Int. J Psychiatry Med., 1989, 19, 2, 203-209.
14. Burke W.J., Nitcher R.L., Rocaforte W.F., Wengel S.P ve ark., The reliability and validity of the Geriatric Depression Rating Scale administered by telephone. JAG 1995. 43: 674-679.
15. Dunn V.K. ve Sacco W.P., Psychometric evaluation of the Geriatric Depression Scale using an elderly community sample, Psychology and Aging 1989, 4, 1, 125-126.
16. Cialdella Ph, Guillaud-Bataille J.M., Gausset M.F. ve ark., Etude sur l'unidimensionnalité de l'échelle de depression gériatrique de Yesavage et Brink. L'Encephale, 1992, XVIII, 537-544.