



## ARAŞTIRMA / RESEARCH

# Klinik araştırmacılar için güvenilirlik ve uyum çalışmalarına ait kontrol listesi GRRAS (Guidelines for Reporting Reliability and Agreement Studies) kılavuzunun Türkçe uyarlaması\*

Turkish adaptation of the guidelines for reporting reliability and agreement studies (GRRAS) for clinical investigators

Semra Erdoğan<sup>1</sup>, Gülhan Orekici Temel<sup>1</sup>

<sup>1</sup>Mersin Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Mersin, Turkey

*Cukurova Medical Journal 2016;41(4):613-621.*

### Abstract

**Purpose:** The primary purpose in this study is to provide to researchers with the Turkish adaptation of GRRAS, which was initially developed for the sake of accurately and completely report reliability and agreement studies. The secondary purpose is to set forth the rate at which GRRAS was utilized in the reliability and agreement related articles published in the scanned journals of the Turkish Medical Index since 2011.

**Material and Methods:** GRRAS was translated into Turkish by two authors. Later on, studies published in the Turkish Medical Index since 2011 were examined in line with the determined criteria.

**Results:** The items included in each section of the Turkish-adapted GRASS were explained. Furthermore, in consequence of the scanning of the Turkish Medical Index, it was determined that of the total 34 studies were published. All these studies were evaluated with the use of the Turkish-adaptation of GRRAS and on the basis of every item included in the guide, reporting percentages were determined by years and in total. Discussion and

**Conclusion:** It was so important for reviewers and editors to make reviews in a standardized way and also for researchers to report their studies in a standard format.

**Key words:** Agreement, reliability, methods comparison, guideline.

## GİRİŞ

Son yıllarda çok sayıda makale ve tez tıp literatürüne eklenmekte, ulusal ve uluslararası bilimsel dergilerde yayınlanmaktadır. Bu süreçte endişe edilen en

### Öz

**Amaç:** Bu çalışmada birincil amaç, güvenilirlik ve uyum çalışmalarının doğru ve tam olarak raporlanabilmesi için geliştirilmiş olan GRRAS kılavuzunun Türkçeye uyarlanmış halini araştırmacılara sunmaktır. İkincil amaç olarak, 2011 yılından itibaren Türk Tıp Dizininde taranan dergilerde yayınlanan güvenilirlik ve uyum ile ilgili makalelerde GRRAS kılavuzundan ne kadar yararlandığını ortaya koymaktır.

**Gereç ve Yöntem:** İki yazar tarafından GRRAS kılavuzu Türkçeye çevrilmiştir. 2011 yılından itibaren Türk Tıp Dizininde yayınlanmış olan çalışmalar belirlenen kriterler doğrultusunda incelenmiştir.

**Bulgular:** Türkçeye uyarlanmış olan GRRAS kılavuzunun her bir bölümünde yer alan maddeler açıklanmıştır. Ayrıca, Türk Tıp Dizininde yapılan tarama sonucunda 34 çalışmanın yayınlanmış olduğu belirlenmiştir. Tüm bu çalışmalar, Türkçe uyarlaması yapılmış GRRAS kılavuzu kullanılarak değerlendirilmiş ve kılavuzda bulunan her bir madde için yıllara ve toplama göre rapor edilme yüzdeleri verilmiştir.

**Sonuç:** Dergi hakem ve editörlerinin makale değerlendirme sürecinde standart bir şekilde değerlendirme yapabilmeleri, ayrıca araştırmacıların da yaptıkları çalışmaları, standart bir formatta raporlayabilmeleri açısından oldukça önemli bir nitelik taşımaktadır.

**Anahtar kelimeler:** Uyum, güvenilirlik, yöntem karşılaştırması, kılavuz.

önemli konu ise yayınlanan çalışmaların kalitesi ve güvenilirliğidir. Klinik araştırmaların kalitesini ve güvenilirliğini artırmak amacıyla pek çok bilimsel dergi, ön değerlendirme aşamasında her bir çalışmaya uygun kılavuz ve standartlardan

Yazışma Adresi/Address for Correspondence: Dr. Semra Erdoğan, Mersin Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Mersin, Turkey. E-mail: semraerdogann@gmail.com  
Geliş tarihi/Received: 03.05.2016 Kabul tarihi/Accepted: 06.06.2016

yararlanmaktadır. Genel tıp dergilerinin editörlerinden oluşan küçük bir grup ilk olarak 1978 yılında, dergilere gönderilecek olan makalelerin biçimi için bir rehber oluşturmak amacıyla bir araya gelmişlerdir. Uluslararası Tıp Dergileri Editörler Komitesi (International Committee of Medical Journal Editors, ICMJE) tarafından Biyomedikal dergilere gönderilen makalelerde bulunması gerekli standartların birçok baskısı gerçekleştirilmiş olup en son 2008 yılında güncellenmiştir<sup>1</sup>. Dolayısıyla bu standartlara uyularak ortak bir yazım dili oluşturulmaktadır. Çalışmanın tasarımına uygun olarak pek çok standart oluşturulmuştur. CONSORT (The Consolidated Standards of Reporting Trials) kılavuzu, Randomize edilmiş Paralel grup çalışmalarının yürütülmesinde ve raporlanmasında kullanılmak üzere geliştirilmiş olan ilk kılavuzlardan birisidir<sup>2</sup>.

CONSORT'tan sonra değişik kılavuzlar da geliştirilmiştir. Randomize kontrollü denemelerin meta analizi için bir kontrol listesi olan QUOROM (Quality of Reporting of Meta-Analyses)<sup>3</sup>, sistematik derleme ve meta analizleri için PRISMA (Preferred Reporting Items for Systematic Reviews and Meta Analyses) kılavuzu en son 2015'de revize edilmiştir<sup>4</sup>. Gözlemsel çalışmaların daha iyi yorumlanabilmesi için geliştirilmiş olan kontrol listesi STROBE (Strengthening the Reporting of Observational Studies in Epidemiology); gözlemsel çalışmaların meta analizleri için ise MOOSE (Proposal for Reporting Meta Analyses of Observational Studies in Epidemiology) kontrol listesi hazırlanmıştır<sup>5</sup>. Tanı doğruluğu çalışmalarının kalitesini artırmak amacıyla QUADAS (Quality Assessment of Diagnostic Accuracy Studies), tanı doğruluğu çalışmalarının sonuçlarının raporlanmasına yönelik kalitenin artırılması amacıyla STARD (Standards for Reporting Studies of Diagnostic Accuracy) kontrol listeleri<sup>6,7</sup>, hayvan deneylerinin daha iyi değerlendirilebilmesi için ARRIVE (Animals in Research: Reporting In Vivo Experiments) kılavuzu önerilmiştir<sup>8</sup>. Olgu sunumu çalışmaları için CARE kontrol listesi (Consensus-based Clinical Case Reporting Guideline Development)<sup>9</sup>, istatistiksel sonuçların yazımı kılavuzu için ise SAMPL kılavuzu (Statistical Analyses and Methods in the Published Literature) önerilmiştir<sup>10</sup>.

Tanı testlerinin güvenilirliği çalışmalarının kalitesini değerlendirmek için 2009 yılında Güvenilirlik çalışmaları için kaliteyi değerlendirme kılavuzu olan

QAREL (The Quality Appraisal for Reliability Studies)<sup>11</sup>, ardından 2010 yılında COSMIN (CONsensus-based Standards for the selection of health Measurement Instruments)<sup>12</sup> ve 2011 yılında da Güvenilirlik ve Uyum çalışmaları için GRRAS (Guidelines for Reporting Reliability and Agreement Studies) kontrol listesi geliştirilmiştir<sup>13</sup>. COSMIN, güvenilirlik ve uyum çalışmalarının metodolojik kalitesi ile ilgilenirken, GRRAS ise bu tür çalışmalar raporlanırken ele alınan metodolojik sorunların tanımlanması ile ilgilenmektedir. QUAREL ise güvenilirlik çalışmalarının sistematik derlemelerinin kalitesini değerlendirmede kullanılan bir kılavuz olarak bu boşluğu doldurmaktadır<sup>14</sup>.

Bu çalışmada amaç, güvenilirlik ve uyum çalışmalarının doğru ve tam olarak raporlanabilmesi için geliştirilmiş olan GRRAS kılavuzunun Türkçeye uyarlanmış halini araştırmacılara sunmaktır. Böylelikle Türkçeye uyarlanmış olan GRRAS kılavuzunun ülkemizde güvenilirlik ve uyum çalışması yapacak olan araştırmacılara bir rehber olması umut edilmektedir. Ayrıca bu çalışmada 2011 yılından itibaren Türk Tıp Dizininde taranan dergilerde yayınlanan güvenilirlik ve uyum ile ilgili makalelerde GRRAS kılavuzundan yararlanmanın ne kadar önemli olduğunu ortaya koymaktır.

## GEREÇ VE YÖNTEM

### Güvenilirlik ve Uyum

Literatürde, "güvenilirlik" ve "uyum" kavramları genellikle birbirlerinin yerine kullanılabilmesine rağmen, kavramsal olarak birbirinden farklılık göstermektedir. Güvenilirlik ve uyum çalışmalarında, istatistik yöntemin ve yorumun doğru yapılabilmesi için güvenilirlik ve uyum kavramları arasındaki farklılığın ortaya konulması gerekmektedir. Güvenilirlik, bir ölçüm sürecinde ölçme işleminin tekrarlanabilirliği ya da tekrarlı ölçümlerin tutarlılığını ifade ederken uyum, bir denekten elde edilen ölçümlerin yakınlığını ölçen bir ölçme aracıdır. Benzer koşullar altında aynı yöntemle aynı denek üzerinden elde edilen tekrarlı ölçüm değerlerinin ölçme hatasından kaynaklı olarak birbirine eşit çıkması beklenen bir durum değildir. Bu nedenle, ölçüm değerleri arasındaki uyumun tespit edilmesi gerekmektedir. Uyum, hem doğruluk hem de kesinlik kavramlarını içermektedir. Doğruluk, ölçüm değerinin gerçek değere yakınlığının bir göstergesi iken kesinlik ise aynı

koşullarda elde edilen tekrarlı ölçümlerin birbirine yakınlığının bir göstergesidir<sup>13,15</sup>.

İki kavramın farklılığını ortaya koymak adına Vet ve arkadaşları (2006) yapmış oldukları çalışmada iki soruya odaklanmak gerektiğini vurgulamaktadırlar. 1. Tekrarlı ölçümler arasındaki **uyum** ne kadar iyi? Yani, elde edilen ölçümler birbirlerine ne kadar benzemektedir. 2. Ölçümler ne kadar **güvenilir**? Diğer bir deyişle, ölçüm hatalarına rağmen, bir hastanın diğerinden farklı olduğu ayırt edilebilir mi? Bu durumda, ölçüm hatası, denekler arasındaki değişkenlikle ilişkilidir<sup>16</sup>.

### GRRAS Kılavuzu

Güvenilirlik ve uyum çalışmalarının sonuçları raporlanırken, çalışma deseninin nasıl yapıldığı, nasıl yürütüldüğü, sonuçların nasıl elde edildiğinin anlaşılabilmesi için yeterli bilgi verilmesi gereklidir. Güvenilirlik ve uyum tahminleri, değerlendiricilerin ya da çalışmaya alınacak olan deneklerin özellikleri, ölçme aracının tipi, istatistiksel yaklaşımlar gibi bazı çeşitli varyasyon kaynaklarından etkilenmektedir. Bu nedenle, çalışma sonuçları yalnızca yorumlanmamalı aynı zamanda ölçme sonuçlarını etkileyecek olan varyasyon kaynakları yeteri kadar tanımlanmalı, hesaplama yöntemleri ve grafiksel gösterimler tamamıyla anlatılmalıdır. Kotner ve arkadaşları (2011) pek çok güvenilirlik ve uyum çalışmalarını gözden geçirildikten sonra, açıkça verilmesi gereken istatistik analizler ve çalışma tasarımı ile ilgili önemli bilgilerin genellikle eksik kaldığını gözlemişlerdir<sup>13</sup>. Raporlamaların yetersiz yapılmasından dolayı çalışma ile ilgili sonuçların sentezlenmesi ve yorumlanması zor olmaktadır. Sağlık hizmetleri ve tıp alanında rapor edilen güvenilirlik ve uyum çalışmalarının planlanması, yürütülmesi ve araştırma sonuçlarının raporlanması aşamalarında bir standardın olmaması, araştırmacıların, yazarların, dergi editörlerinin ve hakemlerinin kullanabileceği resmi bir kılavuz oluşturma fikri ortaya atılmış ve ilk olarak 2009 yılında QUAREL, ardından 2010 yılında COSMIN ve son olarak da 2011 yılında GRRAS kılavuzu geliştirilmiştir<sup>11-13,17,18</sup>. GRRAS kılavuzunun Türkçeye çevrilmiş hali Tablo I'de verilmektedir. Türkçeye uyarlanmış GRRAS kılavuzu 15 maddeden oluşmaktadır. Ancak, güvenilirlik ve tanı testlerinin uyumlarının araştırıldığı çalışmalarda bazen birden fazla test kullanılarak değerlendirmeler yapılmaktadır. Buradan yola çıkılarak, Oster ve arkadaşları (2013), GRRAS kılavuzuna ek olarak bazı

önerilerde bulunmuşlardır. Giriş bölümünde Madde 3'ü, "Verilerin kalitesi ve deneklerin seçiminde kullanılan veri tabanı ifade edilmelidir" şeklinde, yöntem bölümünde Madde 6'nın "Denek ve değerlendirici popülasyonu ve örnekleme yöntemi açıklanmalıdır" şeklinde düzenlenmesini ve "Referans standart tanı testinin belirtilmesi" şeklinde bir madde daha ilave edilmesini önermişlerdir<sup>19</sup>. Böylelikle, çalışmanın kalitesi ve anlaşılabilirliği artıracak gibi çalışma sonuçlarının da doğruluğuna katkı sağlayacaktır.

### Dil Eşdeğerliliği Çalışması

GRRAS Kılavuzunun Türkçeye uyarlanması sürecinde, iki yazar tarafından GRRAS kılavuzu Türkçeye çevrilmiştir. Dil eşdeğerliliğinin sağlanması açısından, Türkçeye çevrilen GRRAS kılavuzu, Yabancı Diller Bölümünden bir öğretim elemanı tarafından kontrol edilerek düzeltmeler yapılmış ve son haline getirilmiştir. Ardından GRRAS kılavuzunun Türkçeye uyarlanmış hali ve kılavuzun orijinal hali Yabancı Diller Bölümünden farklı bir öğretim elemanı tarafından incelenmiş ve karşılaştırıldığında benzer olduğu saptanmıştır.

### Makale Taraması

Türk Tıp Dizininde yapılan tarama sonucunda şu kriterlerin dikkate alınması belirlenmiştir: 1. Makalelerin 2011-2015 yılları arasında yayınlanmış olması, 2. Yayın dilinin Türkçe olması, 3. Başlık, özet ya da anahtar kelimelerde "uyum", "değerlendiriciler arası uyum", "değerlendirici içi uyum", "güvenilirlik", "metot karşılaştırması" kelimelerinin yer almasıdır. Tarama sonucunda 166 makalenin verilen kriterlere uygun olduğu belirlenmiş, makalelerin tam metinleri incelendiğinde ise 34 çalışmanın güvenilirlik ve uyum çalışması olduğu tespit edilmiştir.

### BULGULAR

#### Türkçeye Uyarlanmış GRRAS Kılavuzunun Bölümleri

**Başlık/Özet:** Değerlendiriciler arası güvenilirlik çalışmaları üzerine sistematik derlemelerde, kesitsel veya ileriye yönelik çalışma tasarımlarının gidışı tabyınca elde edilen pek çok güvenilirlik ve uyum tahminleri rapor edilmektedir. Bu araştırmaların çoğu, güvenilirlik ya da uyum gibi anahtar sözcükler kullanılmadığı için kaçınılmazdır. Eğer indekslemede yazılmamışsa güvenilirlik ve uyum

araştırmalarında bunu tespit etmek oldukça zordur. kelimelerin yer alması gerekmektedir<sup>13,20,21</sup>. (Madde Dolayısıyla, başlıkta veya özetle mutlaka bu 1).

**Tablo 1. Güvenilirlik ve Uyum Çalışmalarının Raporlanması için Kullanılan Türkçeye uyarlanmış Kılavuz: GRRAS**

Bölüm ve Konu	Madde	Kotner (2011)	Oster (2013)
Başlık/Özet	1	Başlıkta veya özetle, değerlendiriciler arası/içi güvenilirlik veya uyumun araştırıldığı belirtilmelidir.	
Giriş	2	İlgilenilen tanı testi/ölçme aracı açık bir biçimde tanımlanmalı ve belirtilmelidir.	
	3	İlgilenilen denek popülasyonu açıkça belirtilmelidir.	Verilerin kalitesi ve deneklerin seçiminde kullanılan veri tabanı ifade edilmelidir
	4	Varsa ilgilenilen değerlendirici popülasyonu açıkça belirtilmelidir (eğer uygulanabilirse)	
	5	Güvenilirlik ve uyum kavramları ifade edilmelidir ve çalışma için önemi ve bir gerekçe varsa belirtilmelidir (eğer uygulanabilirse).	
Yöntemler	6	Örneklem büyüklüğünün nasıl seçildiği açıklanmalıdır. Tanımlanan değerlendirici sayısı, denek sayısı ve gözlemlerin tekrar sayıları ifade edilmelidir.	Denek ve değerlendirici popülasyonu ve örnekleme yöntemi açıklanmalıdır.
	7		Referans standart tanı testinin belirtilmesi
	8	Örnekleme yöntemi belirtilmelidir.	
	9	Ölçme/puanlama süreçleri tanımlanmalıdır (tekrarlanan ölçümler arasındaki zaman aralığı, klinik bilgilerin kullanılabilirliği, körleme vb.)	
	10	Ölçme/puanlama süreçlerinin bağımsız yapıp yapılmadığı ifade edilmelidir.	
	11	İstatistiksel analizler belirtilmelidir.	
Sonuçlar	12	Çalışmada yer alan gerçek değerlendirici ve denek sayısı ile gözlemlere ait tekrarlanan ölçüm sayısı ifade edilmelidir.	
	13	Örnekleme alınan değerlendirici ve deneklerin özellikleri belirtilmelidir (eğitim, tecrübe vb.)	
	14	Belirtilmeyen istatistiksel ölçümler dahil güvenilirlik ve uyum istatistiklerinin sonuçları rapor edilmelidir.	
Tartışma	15	Sonuçların pratikle uygunluğu tartışılmalıdır.	
Yardımcı Materyal	16	Eğer mümkünse sonuçların detayları (örneğin online) verilmelidir.	

**Giriş:** Güvenilirlik ve uyum çalışmalarında kullanılacak olan istatistiksel yöntem, sonuç değişkeninin özelliğine yani kullanılan tanı testinin sürekli ya da kategorik olmasına bağlı olarak belirlenmektedir. Dolayısıyla sonuç değişkeninin özelliği (ölçme aracı) açık bir şekilde ifade edilmelidir (Madde 2). Ölçme aracı veya tanı testi belirli yaş grupları, hastalığın belirli evreleri gibi bazı özel popülasyonlar için geliştirilmiş olabilir. Ayrıca, deneklerin özellikleri güvenilirlik ve uyum katsayılarının yorumlanmasını etkilemektedir. Çünkü elde edilen katsayılar bu popülasyondan elde

edilmiştir. Güvenilirlik katsayıları ve uyum istatistikleri belirli bir popülasyona aittir ve çalışılan parametrenin görülme sıklığına bağlıdır. Bu nedenle ilgilenilen popülasyonun özelliği ve bu popülasyondan örneklemin nasıl seçildiği hakkında bilgi verilmelidir (Madde 3)<sup>13,19,22,23</sup>. Ölçekler ya da ölçme araçları çeşitli alanlarda çalışın, çeşitli eğitim seviyelerinden ve bazı özel koşullara sahip kişiler için tasarlanmaktadır. Genellikle, ilgilenilen değerlendirici popülasyonu da söz konusu olan ölçme aracı kullanılarak ve bu alanlarda çalışın tüm kişilerden oluşacaktır. Dolayısıyla, değerlendiricilerin nitelikleri,

linik bilgi düzeyleri, tecrübeleri, uzmanlık derecesi gibi bazı özellikler tanımlanmalıdır (Madde 4)<sup>13,24</sup>. Ayrıca, bu bölümde, güvenilirliğin ve uyumun araştırılmasının neden önemli olduğu ve bu yeni çalışmanın niçin gerekli olduğu ifade edilmelidir (Madde 5)<sup>13</sup>.

**Yöntemler:** Güvenilirlik ve Uyum çalışmalarında deneklerin ve değerlendiricilerin/tanı testlerinin/yöntemlerin nasıl belirlendiği, hangi yöntemle seçildiği, farklı tanı testlerine/değerlendiricilere deneklerin atanma şekli, değerlendirici/tanı testi/yöntem sayısı, referans standart bir testin olup olmadığı açık bir şekilde belirtilmelidir. Çünkü kullanılacak olan teste karar verebilmek için bunların bilinmesi gerekmektedir (Madde 6, Madde 7, Madde 8)<sup>13,20</sup>. Çalışmada ölçümlerin nasıl yapıldığına dair (Madde 9), değerlendiricilerin kararlarının birbirlerinden bağımsız verilip verilmediği konusunda bilgiler verilmelidir (Madde 10). Güvenilirlik ve uyum çalışmalarında kullanılan istatistikler bu bölümde ifade edilmelidir (Madde 11)<sup>13,19,20</sup>.

**Sonuçlar:** Değerlendirici ve denek popülasyonundan, dahil etme ve hariç tutma kriterlerine göre çalışmaya dahil edilen değerlendirici/tanı testi/yöntem, denek ve tekrarlı ölçümler bir akış diyagramı şeklinde verilmelidir (Madde 12). Örnekleme dahil edilen denek ve değerlendiricilere ait özellikler ifade edilmelidir. Çünkü bu özellikteki bir örneklemden elde edilen sonuçların diğer popülasyonlar için genelleme yapıp yapılamayacağı konusunda bize bilgi verecektir (Madde 13). Kullanılan güvenilirlik ve uyum istatistiklerinin sonuçları güven aralıkları ile birlikte verilmelidir (Madde 14)<sup>13</sup>.

**Tartışma:** Güvenilirlik ve uyum istatistikleri için literatürde bazı derecelendirmeler yer almaktadır. Elde edilen yüksek uyum klinikte uyumsuz bir durumu gösterebilir. Sınıf içi korelasyon katsayısında da, minimum uyum için genellikle 0.60, 0.70 ve 0.80 kullanılmaktadır ancak bu durum yalnızca araştırmamızın amacına veya karşılaştırılacak olan grup sayısına göre yeterli olabilir. Mesela, bası yarısı riski için kullanılan ölçüğün klinikte pratik olarak uygulanabilmesi için en azından 0.90 ve üzeri bir değer alması gerekmektedir. Dolayısıyla çalışma sonunda elde edilen güvenilirlik ve uyum sonuçlarını etkileyen faktörler bakımından yorumlanması gerekmektedir (Madde 15)<sup>13</sup>.

**Yardımcı Materyal:** Çalışmada kullanılan ham

veriler ve detaylı sonuçlar, yeniden analiz edilebilmesi ve güvenilirlik ve uyum çalışmalarının meta analizleri ve sistematik derlemeleri yapılabilmesi için çalışmanın sonunda yardımcı materyal olarak verilebilir (Madde 16)<sup>13</sup>.

## Türk Tıp Dizininde yapılan tarama sonuçları

Tarama sonucunda 34 makalenin, 5'inin 2011, 14'ünün 2012, 3'ünün 2013, 11'inin 2014 ve 1 tanesinin de 2015 yılında yayınlanmış olduğu belirlenmiştir. Tüm bu çalışmalar, Türkçe uyarlaması yapılmış GRRAS kılavuzu kullanılarak değerlendirilmiş ve kılavuzda bulunan her bir madde için yıllara ve toplama göre rapor edilme yüzdeleri Tablo 2'de verilmiştir.

Tablo 2 incelendiğinde, GRRAS kılavuzunun ilk geliştirildiği yıl olan 2011 yılında, Madde 1, 6, 11, 14 ve 15'in tüm çalışmalarda raporlandığı, 2012 yılından itibaren bu maddelere ilaveten Madde 2'nin de raporlanmaya başlandığı gözlenmiştir. 2011 yılında, Madde 4, 8, 10, 12 ve 13'ün hiç rapor edilmediği, Madde 10, 12 ve 13'ün ise diğer yıllarda da neredeyse hiç rapor edilmediği gözlenmiştir. Toplamda makalelerin en az % 80'inde raporlanan 6 kriter (Madde 1,2,6,11,14 ve 16), en fazla % 10'unda raporlanan 6 kriter (Madde 3,4,5,8,12,13), % 10 - % 45 arasında 3 kriter (Madde 7, 9 ve 10) bulunmaktadır.

## Sınırlılıklar

Bu çalışma ile güvenilirlik ve uyum çalışmalarında kullanılan bir kontrol listesini Türkçe'ye uyarlama çalışması yapılmıştır. Çalışmada sadece ilgi alanı güvenilirlik ve uyum olan iki Biyoistatistik uzmanı tarafından gerçekleştirilmiştir. Güvenilirlik ve uyum çalışmalarının Türkiye tabanlı değerlendirilmesi aşamasında sadece Türk Tıp Dizininde yapılan tarama sonucunda ele alınmıştır. Çalışmalar seçilirken anahtar kelimelerde veri tabanının kullanım kısıtlılıklardan kaynaklı olarak sadece "uyum", "değerlendiriciler arası uyum", "değerlendirici içi uyum", "güvenilirlik" ve "metot karşılaştırması" kelimeleri yer almıştır.

## TARTIŞMA

Güvenilirlik ve uyum kavramsal olarak birbirinden farklı olmasına rağmen genellikle karıştırılan ve birbiri yerine kullanılan kavramlardır. Güvenilirlik, hastalar/denekler/değerlendiriciler arası farklılığı ortaya koyarken, uyum zaman içerisindeki

değişimi/etkileri değerlendirilmek için kullanılmaktadır.

**Tablo 2. 2011-2015 yılları arasında Türk Tıp Dizininde yayınlanmış olan çalışmaların GRRAS kriterlerine göre rapor edilme yüzdeleri**

Bölüm ve Konu	Madde	2011 (n=5)	2012 (n=14)	2013 (n=3)	2014 (n=11)	2015 (n=1)	Toplam (n=34)
Başlık/Özet	1	4 (%80)	13 (%93)	3 (%100)	11 (%100)	1 (%100)	32 (% 94)
Giriş	2	2 (%40)	14 (%100)	3 (%100)	9 (%82)	1 (%100)	29 (%85)
	3	1 (%20)	0	0	0	1 (%100)	2 (% 6)
	4	0	0	0	0	0	0
	5	2 (%40)	1 (% 7)	0	0	0	3 (%9)
	6	5 (%100)	14 (% 100)	3 (%100)	11 (%100)	1 (%100)	<b>34 (%100)</b>
Yöntemler	7	3 (%60)	6 (%43)	1 (%33)	4 (%36)	1 (%100)	15 (%44)
	8	0	0	0	1 (%9)	0	1 (%3)
	9	3 (%60)	3 (%21)	2 (%67)	5 (%45)	0	13 (% 38)
	10	0	2 (% 14)	1 (%33)	2 (%18)	0	5 (%15)
	11	5 (%100)	12 (%86)	3 (%100)	9 (%82)	1 (%100)	30 (% 88)
	12	0	0	0	2 (%18)	0	2 (% 6)
Sonuçlar	13	0	0	0	1 (%9)	0	1 (% 3)
	14	5 (%100)	13 (%93)	3 (%100)	9 (%82)	1 (%100)	31 (% 91)
Tartışma	15	5 (%100)	14 (% 100)	3 (%100)	11 (%100)	1 (%100)	<b>34 (% 100)</b>
Yardımcı Materyal	16	0	0	0	0	0	0

Güvenilirlik ve uyum çalışmalarında, uygulanacak istatistiksel yöntem, ölçüm sonucunun (sonuç değişkeninin) nasıl bir değişkenle ifade edildiğine yani kullanılan tanı testinin sürekli ya da kategorik olması, değerlendirici sayısına bağlı olarak belirlenmektedir. Sağlık alanında yapılan çalışmalarda da sonuç değişkeni, çekilmiş olan röntgen filminden/diğer görüntüleme tekniklerinden elde edilen sonuçlar, biyopsi sonuçları, laboratuvar testleri ve fiziksel tetkiklerden elde edilen bulgular şeklinde olabilir. Özellikle biyokimya, patoloji, radyoloji, mikrobiyoloji vb. alanlarda tek bir değerlendirici tarafından alınan ölçüm sonuçlarının uyumu/güvenilirliği ya da birden fazla değerlendirici tarafından alınan ölçüm sonuçlarının uyumu/güvenilirliği araştırılmak istenebilir. Ayrıca, klinisyenler, herhangi bir hastalığın tanısının konulmasında, yaygın olarak kullanılan klasik yöntemlere göre düşük maliyetli ve tedaviye daha kısa bir sürede cevap veren ya da tanının hızlı ve doğru konulduğu bir yöntem bulduklarında, bu yöntemi referans (eski, klasik) yöntem ile karşılaştırarak yöntemlerden elde edilen sonuçların ne derece uyumlu olduklarını belirlemek isteyebilirler<sup>14,15</sup>.

Örneğin, tanı ve tedavi yöntemlerindeki ilerlemelere rağmen, Tüberküloz hastalığının önlenmesi, tanısının konulması ve tedavisinin hızlı ve doğru bir şekilde yapılabilmesi için yöntem arayışları

sürmektedir. Günümüzde mikobakterilerin tanımlanmasında konvansiyonel kültür yöntemleri rutin mikobakteriyoloji laboratuvarlarında hala altın standart (gold standart) olarak kabul edilmektedir. Ancak, bu mikroorganizmaların yavaş üremesi, kültür yöntemlerinin zor/zahmetli olması ve uzun sürmesi gibi dezavantajları bulunmaktadır.

Mikobakterilerin daha kısa sürede, kolay bir şekilde ve daha düşük maliyetli yöntemlerle tespit edilmesi oldukça önemli bir durumdur. Dolayısıyla, güvenilirlik ve performans açısından geleneksel yöntemlere göre daha avantajlı olan farklı moleküler teknikler, rutin mikobakteriyoloji laboratuvarlarında uygulanabilirlikleri açısından karşılaştırılarak yeni yöntemler önerilebilir<sup>25</sup>.

Bu tür çalışmalarda, çalışmalarının planlanması, yürütülmesi ve araştırma sonuçlarının raporlanması oldukça önemlidir. Sağlık hizmetleri ve tıp alanında rapor edilen güvenilirlik ve uyum çalışmaları için çok yaygın kabul edilen resmi kriterler, standartlar ya da kılavuzlar bulunmamaktadır<sup>17</sup>. Sağlık alanındaki bu tür çalışmalarda, bir standardın olmadığından, araştırmacıların, yazarların, dergi editörlerinin ve hakemlerinin kullanabileceği resmi bir kılavuz oluşturma fikri gelişmiş ve Kotner ve arkadaşları (2011) tarafından GRRAS kılavuzu geliştirilmiştir<sup>13</sup>. Bazen güvenilirlik ve tanı testlerinin uyumlarının

araştırıldığı çalışmalarda bazen birden fazla test kullanılarak değerlendirilmeler yapılmaktadır. Buradan yola çıkılarak, Oster ve arkadaşları (2013), GRRAS kılavuzuna ek olarak bazı önerilerde bulunmuşlardır<sup>19</sup>.

**Tablo 3. IMRAD Formatına göre Türkçeye Uyarlanmış GRRAS kılavuzu**

Bölüm ve Konu	Madde	
Başlık/Özet	1	Başlıkta veya özetle, değerlendiriciler arası/içi güvenilirlik veya uyumun araştırıldığı belirtilmelidir.
Giriş	2	Güvenilirlik ve uyum kavramları açıklanmalı, çalışma için gerekliliğinden ve öneminden bahsedilmelidir.
	3	İlgilenilen tanı testi/ölçme aracı ve referans standart tanı testi açık bir biçimde açıklanmalıdır.
Yöntemler	4	İlgilenilen denek ve değerlendirici popülasyonu belirtilmelidir.
	5	Örneklem büyüklüğünün nasıl belirlendiği ve hangi yönetime göre seçildiği açıklanmalıdır.
	6	Tanımlanan değerlendirici sayısı, denek sayısı ve gözlemlerin tekrar sayısı, tekrarlanan ölçümler arasındaki zaman aralığı ifade edilmelidir.
	7	Değerlendirici ve deneklerin özellikleri (eğitim, tecrübe vb.), ölçümlerin birbirinden bağımsız yapıp yapılmadığı belirtilmelidir.
	8	İstatistiksel analizler belirtilmelidir.
Sonuçlar	9	Belirtilmeyen istatistiksel ölçümler dahil güvenilirlik ve uyum istatistiklerinin sonuçları rapor edilmelidir.
Tartışma	10	Elde edilen sonuçların pratikle uygunluğu tartışılmalıdır.
Yardımcı Materyal	11	Eğer mümkünse sonuçların detayları (örneğin online) verilmelidir.

Türk Tıp Dizininde yapılan tarama sonucunda, GRRAS kılavuzunda bulunan maddelerin buldukları bölümlerin, makalelerde araştırmacıların farklı bölümlerde yer verdiği gözlenmiştir. Giriş bölümünde, güvenilirlik/uyum hakkında bilinmesi gereken ifadesi içeren Madde 5, Giriş bölümünde değil yöntemler bölümünde açıklandığı pek çok makaleye rastlanmıştır. Bunun yanı sıra, sonuçlar bölümünde yer alan "çalışmada yer alan değerlendirici, denek sayısı, gözlemlere ait tekrarlanan ölçüm sayılarının, değerlendirici ve deneklerin özelliklerini ifade eden" maddelerin (Madde 12 ve Madde 13) araştırmacılar tarafından yöntemler bölümünde ele alındığı gözlenmiştir. Ayrıca, incelenen 34 makalenin hiç birinde bir akış diyagramına rastlanmamıştır (Tablo 2). Genel olarak bilimsel bir yazı, Giriş (Introduction), Gereç ve Yöntem (Material and Method), Bulgular (Sonuçlar, Results) ve (And) Tartışma (Discussion) ana bölümlerinin baş harflerinden (IMRAD) oluşan bir yazım formatına göre düzenlenmektedir. Ülkemizde de genel olarak, yapılan çalışmalar IMRAD yazım kurallarına uygun bir şekilde düzenlenmektedir. IMRAD formatında yazılacak olan bilimsel yazının bölümlerinde yer alması gereken ifadelerin ve açıklamaların da bir formatı bulunmaktadır<sup>26</sup>.

Güvenilirlik ve Uyum çalışmalarında kullanılan Kottner (2011) tarafından geliştirilmiş ve Oster (2013) tarafından güncellenmiş olan GRRAS kılavuzunun bazı maddelerin bazı bölümlerde tekrarlandığı ve bazı maddelerin farklı bölümlerde yer aldığı gözlenmiştir. Buradan yola çıkılarak, IMRAD formatına uygun bir şekilde, GRRAS kılavuzu Tablo III'de yeniden düzenlenmiş olup araştırmacıların bu kılavuzu kullanmaları tavsiye edilmektedir.

Sonuç olarak, Türkçe uyarlaması yapılmış GRRAS kılavuzunun, güvenilirlik ve uyum çalışmalarının, planlama, tasarım, yöntem ve verilerin raporlanması bölümlerinin uluslararası kılavuzlara ve standartlara uygun bir şekilde oluşturulması sağlanmıştır. Aynı zamanda, dergi hakem ve editörlerinin makale değerlendirme sürecinde standart bir şekilde değerlendirme yapabilmeleri ve araştırmacılar için de yaptıkları çalışmaları, standart bir formatta raporlayabilmeleri açısından oldukça önemli bir nitelik taşımaktadır. Bunun yanı sıra bu çalışmanın, ülkemizde yapılacak olan güvenilirlik ve uyum çalışmalarında, araştırmacılar, dergi hakem ve editörleri için yardımcı bir kılavuz olması hedeflenmiştir.





**\*Bu çalışmanın bir bölümü 17. Ulusal Biyoistatistik Kongresinde Poster Bildiri olarak sunulmuştur.**

## KAYNAKLAR

1. Şimşek G, Alicura S, Yılmaz O. Biyomedikal Dergilere Gönderilen Makalelerde Bulunması Gerekli Standartlar: Biyomedikal Yayınların Yazımı ve Değerlendirilmesi. *Epilepsi*. 2010;16:133-40.
2. Sunay D, Şengezer T, Oral M, Aktürk Z, Shulz KF, Altman DG ve ark. CONSORT 2010 Raporu: Randomize Paralel Grup Çalışmalarının Raporlanmasında Güncellenmiş Kılavuzlar. *Avrasya Aile Hekimliği Dergisi*. 2013;2:1-10.
3. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*. 1999;354:1896-900.
4. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *The BMJ*. 2009;339:b2535.
5. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson D, Rennie D et al. Meta-analysis of Observational Studies in Epidemiology. A Proposal for Reporting. *JAMA*. 2000;283:2008-12.
6. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. Standards for Reporting of Diagnostic Accuracy. *Clin Chem*. 2003;49:1-6.
7. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.
8. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLOS Biology* 2010;8:e1000412.
9. Gagnier JJ, Kienle G, Altman DA, Moher D, Sox H, Riley D et al. The CARE Guidelines: consensus-based clinical case reporting guideline development. *Journal of Medical Case Reports*. 2013;7:223.
10. Lang TA, Altman DG. Basic statistical reporting for articles published in clinical medical journals: the "Statistical Analysis and Methods in the Published Literature" or the SAMPL Guidelines. *International Journal of Nursing Studies*. 2015;52:5-9.
11. Lucas NP, Macaskill PM, Irwig L, Bogduk N. The Development of a quality appraisal tool for studies of diagnostic reliability (QUAREL). *J Clin Epidemiol*. 2010;63:854-61.
12. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Startford PW, Knol DL et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19:539-49.
13. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64:96-106.
14. Lucas N, Macaskill P, Irwig L, Moran R, Rickards L, Turner R et al. The reliability of a quality appraisal tool for studies of diagnostic reliability (QUAREL). *BMC Med Res Methodol*. 2013;13:111.
15. Barnhart H, Haber B, Lin L. An overview on assessing agreement with continuous measurement. *J. Biopharm Stat*. 2007;17:529-69.
16. Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59:1033-39.
17. Audige L, Bhandari M, Kellam J. How Reliability are reliability studies of fracture classifications? *Acta Orthop Scand*. 2004;75:184-94.
18. Brorson S, Hróbjartsson A. Training improves agreement among doctors using the Neer system for proximal humeral fractures in systematic review. *J. Clin Epidemiol*. 2008;61:7-16.
19. Oster NV, Carney PA, Allison KH, et al. Development of a diagnostic test set to assess agreement in breast pathology: practical application of the Guidelines for Reporting Reliability and Agreement Studies (GRRAS). *BMC Women's Health*. 2013;13:1-8.
20. Hernaez R. Reliability and agreement studies: a guide for clinical investigators. *The BMJ*. 2015;64:1018-27.
21. Kottner J, Raeder K, Halfens R, Dassen T. A Systematic review of interrater reliability of pressure ulcer classification system. *J Clin Nurs*. 2009;18:315-36.
22. Dunn G. *Methods comparison 1: paired observation. Statistical Evaluation of Measurement Errors: Design and Analysis of Reliability Studies*, 2nd ed, London, Hodder Arnold, 2004.
23. Gjørup T. Kappa coefficient and the prevalence of a diagnosis. *Methods Inf Med*. 1988;27:184-6.
24. Nanda R, Gupta S, Kanapathipillai P, Liow RYL, Rangan A. An assessment of the inter examiner reliability of clinical tests for subacromial impingement and rotator cuff integrity. *Eur J Orthop Surg Traumatol*. 2008;18:495-500.
25. Börekci G, Aslan G, Aydın E, Fiandaca MJ, Stender H, Lee NM et al. Identification of Mycobacterium species from BACTEC MGIT™ positive cultures with Oligo-FISH and PNA-FISH methods. *Mikrobiyol Bul*. 2014;48:385-401.
26. Özdamar K. *Bilimsel Yazı Yazım Kuralları. Modern Bilimsel Araştırma Yöntemleri*, 1st ed, Eskişehir, Kaan Kitabevi, 2003.