

Sağlık Araştırmalarında Sınıf İçi Korelasyon Katsayısının Kullanımı

The Use of Intraclass Correlation Coefficient (ICC) in Medical Research: Review

Can ATEŞ,^a
Derya ÖZTUNA,^a
Yasemin GENÇ^a

^aBiyostatistik AD,
Ankara Üniversitesi Tıp Fakültesi,
Ankara

Geliş Tarihi/Received: 25.05.2009
Kabul Tarihi/Accepted: 17.07.2009

Yazışma Adresi/Correspondence:
Derya ÖZTUNA
Ankara Üniversitesi Tıp Fakültesi,
Biyostatistik AD, Ankara,
TÜRKİYE/TURKEY
dgokmen2001@yahoo.com

ÖZET Güvenirlik, ölçümlerin tekrarlanabilirliği ya da tekrarlı ölçümlerin tutarlılığı olarak tanımlanır. Sağlık alanında en yaygın kullanılan güvenilirlik çalışmaları, gözlemci-İçi ve gözlemciler arası uyumdur. Bu çalışmalarda en temel amaç, aynı denekten alınan tekrarlı ölçümler veya aynı denek üzerinde iki ya da daha fazla sayıda gözlemcinin ölçümleri arasındaki uyumun değerlendirilmesidir. Ölçümler sürekli olduğunda, gözlemci İçi veya gözlemciler arası uyumun değerlendirilmesinde, sınıf İçi korelasyon katsayısı (SKK) kullanılır. Çalışmanın amacı, tasarımı ve ölçümlerin alınma tipine bağlı olarak SKK hesaplanmasında kullanılan farklı formüller mevcuttur. Sadece deneklerden kaynaklanan sistematik bir değişkenlik varsa, tek yönlü modeller; deneklerin yanısıra, gözlemcilerden kaynaklanan bir değişkenlik de varsa iki yönlü modeller kullanılmalıdır. Tüm modeller için denek etkisi rastgele alınır. Bununla beraber gözlemcinin etkisi, modelin tanımlanmasında etkilidir. Eğer gözlemcinin etkisi rastgele ise, model "rastgele etki modeli", sabit ise, "karma etki modeli" olarak tanımlanır. Model seçiminden sonra, diğer iki önemli ayırım ise, SKK'nın, elde edilen gözlem değerlerinin tek bir ölçüm sonucuna mı yoksa birden çok ölçümün toplamı ya da ortalaması kullanılarak mı hesaplandığı ve SKK hesaplamalarının, "tutarlılık" veya "mutlak uyum" tanımlarından hangisi kullanılarak yapıldığıdır. SKK, sağlık alanındaki çalışmalarda sıkça kullanıldığı için, hangi deney düzeninde hangi SKK'nın uygun olduğuna karar vermek önemlidir. Bu çalışmada, SKK modelleri arasındaki farklılıklar ayrıntılı olarak açıklanacaktır.

Anahtar Kelimeler: Sınıf İçi korelasyon katsayısı, güvenilirlik, gözlemci İçi/arası uyum

ABSTRACT Reliability is defined as the repeatability of measurements or the consistency of repeated measurements. The most common type of reliability in medical research is the intra/inter-rater reliability. The main purpose of these studies is to evaluate the agreement between repeated measurements from the same subject or the measurements of two or more different raters on the same subject. When the measurements are continuous, intraclass correlation coefficient (ICC) is used to evaluate the intra/inter-rater reliability. There are different formulas for calculating the ICC which depend on the purpose of the study, the design of the study and type of measurements taken. If there is a systematic variability due to subjects only, the one way models should be used. If there is also another systematic variability due to raters in addition to subjects, two way models should be used. For all models, the effect of subjects is always random. However, the effect of raters has an impact on these models. If the effect of raters is random, the models are named as "random effect models", whereas the effect of raters is fixed, those are called as "mixed effect models". After the selection of model, the other important distinctions are to determine whether the measurements used are single or average of two or more measurements taken by different raters and whether the ICC measures the agreement using consistency or absolute agreement definition. Although ICCs are commonly used in medical studies, it is not known which ICC is appropriate for which experimental design. In this paper, the distinctions between different ICCs will be introduced in detail.

Key Words: Intraclass correlation coefficient, reliability, inter/intra rater reliability

Güvenirlilik, bir ölçüm sürecinde ölçüm işleminin tekrarlanabilirliği ya da tekrarlar da ki tutarlılık olarak tanımlanabilir. Yani bir ölçümün güvenilir olmasının koşulu, aynı gözlemciyle/ölçüm aracıyla, aynı koşullar altında, aynı denekten, aynı yanıtın alınmasıdır.¹ Sağlık bilimlerinde, gözlemci uyumunun incelendiği güvenilirlik çalışmalarına sıklıkla rastlanmaktadır. Bu tür çalışmalarda en temel amaç, aynı denekten elde edilen çoklu ölçümlerin uyumunun değerlendirilmesidir. Bir denekten elde edilen çoklu ölçüm, aynı gözlemcinin tekrarlı ölçümleri olabileceği gibi, iki ya da daha fazla sayıda gözlemcinin ölçümleri de olabilir. Ölçümler sürekli olduğunda ilk durumda, gözlemci-içi (intra-rater), ikinci durumda ise gözlemciler-arası (inter-rater) uyum, sınıf içi korelasyon katsayısı (SKK) ya da uyum ilişki katsayısı kullanılarak değerlendirilir.^{2,3}

SKK, hem birim hem de varyansları bakımından aynı sınıfa ait ölçümler arasındaki ilişki miktarını belirlemek amacıyla geliştirilmiştir. Örneğin iki radyoloğun panoromik röntgenlerde alt çene kortikal kemik kalınlığı (mm) ölçümleri arasındaki ilişki, bir batında doğan yavruların ağırlıkları (g) arasındaki ilişki, ikizlerin IQ değerleri arasındaki ilişkinin ölçümünde SKK kullanılır.⁴

Güvenirlilik çalışmalarında SKK'nın hesaplanmasında n denekten k sayıda ölçüm alındığı ve deneklerin oluşturduğu örneklemin büyük bir kitleden rastgele seçildiği varsayılır. Elde edilen ölçümler aynı gözlemcinin tekrarlı ölçümleri olabileceği gibi, farklı gözlemcilerden elde edilen ölçümler de olabilir (Tablo 1).

TABLO 1: SKK hesaplamasında kullanılan veri seti.

Denek (i)	Ölçüm/Gözlemci (j)			
	1	2	...	k
1	X ₁₁	X ₁₂	...	X _{1k}
2	X ₂₁	X ₂₂	...	X _{2k}
...
n	X _{n1}	X _{n2}	...	X _{nk}

UYGUN SINIFIÇİ KORELASYON KATSAYISININ SEÇİMİ

SKK, aynı denekten elde edilen ölçümlerin değişkenliğini, denekler ve ölçümler/gözlemciler üzerinden elde edilen toplam değişkenlik ile karşılaştırır. Bu açıdan değerlendirildiğinde, SKK en genel biçimiyle iki varyansın birbirine oranıdır:

$$SKK = \frac{\sigma_{DA}^2}{\sigma_{DA}^2 + \sigma_{DI}^2} \quad (1)$$

Bu eşitlikte kullanılan, σ_{DA}^2 ; denekler arası varyansı, σ_{DI}^2 ; denekler içi varyansı ifade etmektedir ve bu değerler varyans analizinden elde edilen kareler ortalamaları kullanılarak hesaplanır.⁵

Çalışmanın tasarımına (tek yönlü rastgele etki modeli, iki yönlü rastgele etki modeli, iki yönlü karma etki modeli), amacına (tutarlılık, mutlak uyum) ve kullanılan ölçümlerin elde edilme biçimine (tek ölçüm, toplam/ortalama ölçüm) bağlı olarak farklı SKK tahminleri bulunmaktadır. Farklı SKK tahminlerinin elde edilmesinde belirleyici olan faktörler aşağıda sırasıyla incelenecektir.

ÇALIŞMANIN TASARIMINA GÖRE SKK TAHMİNLERİ

En uygun SKK'nın seçimi için verilmesi gereken ilk karar, veri setinin tek yönlü ya da iki yönlü varyans analizi modellerinden hangisi kullanılarak değerlendirileceğidir. SKK hesaplanacak tüm durumlarda, sistematik varyans kaynaklarından birisi, denekler arasındaki farklılık, ikincisi ise ölçümler/gözlemciler arasındaki farklılıktır. Denek faktörü, varyans analizi modellerinde her zaman rastgele faktör olarak değerlendirilirken; ölçüm/gözlemci etkisi rastgele ya da sabit olabilir.⁶

RASTGELE VE SABİT ETKİ

Teoride, ilgilenilen faktörün düzeylerinin büyük bir kitleden rastgele seçildiği ve o kitleyi temsil ettiği düşünülürse, bu faktörün etkisi rastgele olarak tanımlanır. Örneğin iki radyoloğun 300 dpi çözünürlükte dijitize edilen filmlerden ölçtükleri yükseklikler arasındaki ilişki incelenmek istensin. Bu örnekte rastgele etki modelinin uygulanabilmesi için ölçümü yapacak iki radyoloğun tüm radyolog-

lar içinden rastgele seçildiği, yani ölçümü yapan bu iki radyolog değiştirilse de sonucun değişmeyeceği varsayılır.

İlgilenilen faktörün düzeyleri, araştırma sorusuna bağlı olarak belirleniyorsa ve seçilen düzeylerin değiştirilmesi, araştırmacının hipotezini tamamen farklılaştırıyorsa sabit etkiden bahsedilir. Aynı deneklerden alınan konvansiyonel diş filmlerinden 300dpi ve 600dpi çözünürlükte elde edilmiş görüntüler üzerinde yapılan ölçümlerin ilişkisinin hesaplanmasında ölçümlerin etkisinin sabit olduğu düşünülmelidir.^{4,7}

SKK hesaplamasında kullanılan tek yönlü modellerde ölçüm/gözlemci etkisi her zaman rastgele alınır ve kullanılan model “tek yönlü rastgele etki modeli” olarak adlandırılır. İki yönlü modellerde ölçüm/gözlemci etkisinin rastgele ya da sabit olmasına bağlı olarak, kullanılan modeller sırasıyla “iki yönlü rastgele etki modeli” ve “iki yönlü karma etki modeli” olarak adlandırılır.

TEK YÖNLÜ RASTGELE ETKİ MODELİ

Tek yönlü rastgele etki modelinde, her n denek için elde edilen k sayıda ölçümün, birbirinden farklı nxk sayıda gözlemci tarafından yapıldığı varsayılır ve varyans analizi modeli aşağıdaki gibi tanımlanır:

$$X_{ij} = \mu + d_i + h_{ij} \quad i=1,2,\dots,n; j=1,2,\dots,k \quad (2)$$

Bu eşitlikte, μ , tüm ölçümler için kitle ortalamasını, d_i denek etkisini, h_{ij} ise hata terimini gösterir. Bu modelde, denek etkisinin rastgele ve bağımsız olduğu, $N(0, \sigma_d^2)$ ile normal dağılım gösterdiği, hata teriminin ise rastgele ve bağımsız olduğu, $N(0, \sigma_h^2)$ ile normal dağılım gösterdiği varsayılır.

Bu modelde ölçümleri yapan gözlemci sayısı, tekrarlı ölçüm sayısından daha büyüktür. Örneğin, 10 hastanın, MR merkezleri kitlesinden rastgele seçilen merkezler tarafından değerlendirileceği bir çalışma tasarlandığında ve her bir hastadan iki ölçüm alındığında, bu çalışmada toplam 20 merkez yer alacak ve her bir hastadan alınan iki ölçüm arasındaki uyumun incelenmesinde tek yönlü rastgele etki modelini temel alan SKK hesaplanacaktır.

Yukarıda verilen örnek, tek yönlü rastgele etki modeline uygun olmasına rağmen, uyum çalışmalarında nadir karşılaşılan bir deney düzenidir. Tek yönlü rastgele etki modeli, daha yaygın olarak hangi ölçümün birinci, hangi ölçümün ikinci sırada yer alacağına belirlenemediği durumlarda kullanılır. Örneğin, ikizlerin boy uzunlukları arasındaki ilişki incelendiğinde, hangi ölçümün birinci, hangisinin ikinci ikize ait olduğu arasında ayırmaya yapmak mümkün değildir. Bu örnekte, ikizlerin farklı sırada analize alınması hesaplanan SKK değerlerini değiştirmez. Fakat sıklıkla yapılan bir hata, bu örnek için Pearson korelasyon katsayısının hesaplanmasıdır. Bu durumda ölçüm çiftlerinin yerlerinin değişmesi, farklı Pearson korelasyon katsayısı değerlerinin elde edilmesine neden olur.

İKİ YÖNLÜ RASTGELE ETKİ MODELİ

n sayıda denek için büyük bir kitleden rastgele seçilen k sayıda gözlemci tarafından değerlendirildiği iki yönlü rastgele etki modelinde, varyans analizi modeli aşağıdaki gibi tanımlanır:

$$X_{ij} = \mu + d_i + g_j + h_{ij} \quad i=1,2,\dots,n; j=1,2,\dots,k \quad (3)$$

Bu eşitlikte, μ , tüm ölçümler için kitle ortalamasını, d_i denek etkisini, g_j gözlemci etkisini, h_{ij} ise hata terimini gösterir. Bu modelde, denek, gözlemci etkisi ve hata teriminin rastgele ve bağımsız olduğu, sırasıyla $N(0, \sigma_d^2)$, $N(0, \sigma_g^2)$ ve $N(0, \sigma_h^2)$ ile normal dağılım gösterdikleri varsayılır.

Örneğin, ilgili kitleden rastgele seçilen dört radyologun, hastaları 10’lu Likert ölçeğine göre değerlendirdiği bir durumda, radyologlar arasındaki uyum, iki yönlü rastgele etki modeli kullanılarak hesaplanan SKK ile incelenmelidir.

İKİ YÖNLÜ KARMA ETKİ MODELİ

İki yönlü karma etki modelinde, n sayıda denek k sayıda gözlemci tarafından değerlendirilir ve varyans analizi modeli aşağıdaki gibi tanımlanır. Bu modelde gözlemci etkisinin sabit olduğu varsayılır:

$$X_{ij} = \mu + d_i + g_j + h_{ij} \quad i=1,2,\dots,n; j=1,2,\dots,k \quad (4)$$

Bu eşitlikte, μ , tüm ölçümler için kitle ortalamasını, d_i denek etkisini, g_j gözlemci etkisini, h_{ij} ise hata terimini gösterir. Bu modelde, denek, göz-

lemci etkisi ve hata teriminin rastgele ve bağımsız olduğu, sırasıyla $N(0, \sigma_d^2)$, $N(0, \theta_g^2)$ ve $N(0, \sigma_h^2)$ ile normal dağılım gösterdikleri varsayılır. θ_g^2 , iki yönlü rastgele etki modelindeki σ_g^2 'ye karşılık gelir ve $\theta_g^2 = \sum_{j=1}^k g_j^2 / (k-1)$ eşitliği ile tanımlanır.

Örneğin hastaların iki ayrı uzmanlık seviyesinden iki radyolog tarafından 10'lu Likert ölçeğine göre değerlendirildiği bir çalışmada radyologlar arası uyumun SKK ile incelenmesinde iki yönlü karma etki modeli kullanılır.

İki yönlü rastgele etki ve iki yönlü karma etki modelleri için denek*gözlemci etkileşimini içeren modeller de bulunmaktadır. Fakat çalışmamızda bu modeller ele alınmamıştır. En genel kapsamıyla bu üç model için varyans analizi tablosu ve varyans bileşenleri cinsinden SKK tahminleri Tablo 2'de sunulmuştur. Bu tabloda yer alan σ_d^2 ; deneklerden kaynaklanan varyansı, σ_g^2 ve θ_g^2 ; (sırasıyla gözlemci etkisinin rastgele ve sabit olduğu durumda) gözlemcilerden kaynaklanan varyansı σ_h^2 ; hata varyansını göstermektedir.^{4,5}

ÇALIŞMANIN AMACINA GÖRE SKK TAHMİNLERİ

SKK hesaplamaları, ilişkinin “tutarlılık” ya da “mutlak uyum” olarak değerlendirilmesine bağlı

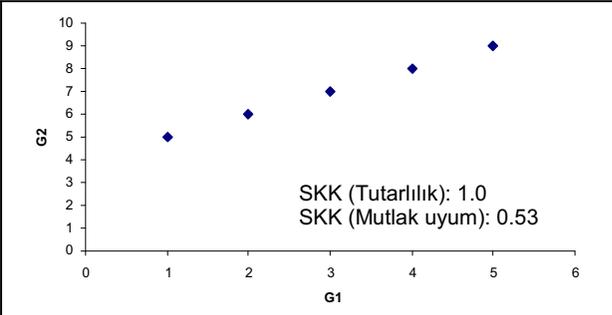
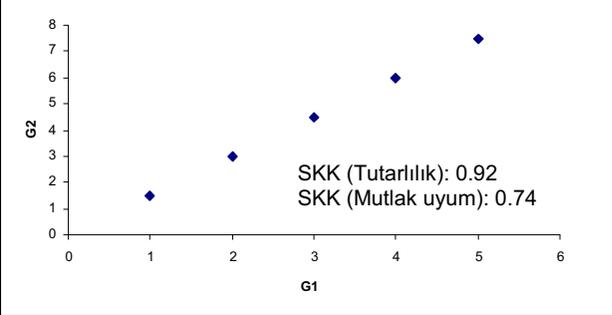
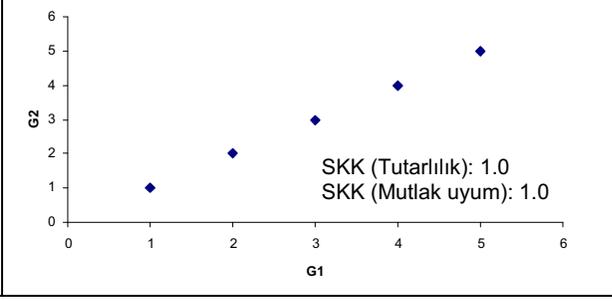
olarak değişir. Tek yönlü modellerde bu tür bir ayırım yokken, iki yönlü modellerde hem tutarlılığı hem de mutlak uyumu temel alan SKK tahminleri elde edilebilir. Tutarlılık ve mutlak uyum arasındaki farklılık, ölçümlerden ya da gözlemcilerden kaynaklanan sistematik değişkenliğin nasıl değerlendirildiği ile açıklanabilir. Eğer değişkenlik önemsiz ise, SKK formülünün payda kısmında bu değişkenliğe ilişkin varyans terimi yer almaz ve elde edilen SKK, “tutarlılık” ölçümü olarak değerlendirilir. Ölçümler ya da gözlemciler arasındaki sistematik değişkenlik önemli ise, ölçüm/gözlemci değişkenliği SKK formülünün paydasında yer alır ve hesaplanan SKK, “mutlak uyum” olarak değerlendirilir.⁸

Mutlak uyumun değerlendirildiği çalışmalarda, istenen durum aynı gözlemci tarafından elde edilen çoklu ölçümlerin ya da farklı gözlemciler tarafından yapılan ölçümlerin birebir aynı olmasıdır. Tutarlılığın değerlendirildiği çalışmalarda ise elde edilen SKK tahminleri toplanabilirlik indeksidir ve bu indeks bir değişkeninin (y) bir diğer değişkene (x) bir sabit ile ($y=x+b$) eşitlenmesinin derecesini ölçer.⁴ Tutarlılık ve mutlak uyum tanımlarına göre SKK tahminlerinin nasıl değişkenlik gösterdiği Tablo 3'te verilen örneklerle sunulmuştur.

TABLO 2: Varyans analizi modelleri için varyans bileşenleri ve SKK tahminleri.

Model ve değişim kaynağı	Serbestlik Derecesi	Kareler Ortalaması	Beklenen Kareler Ortalaması	SKK Tahmini
Tek Yönlü Rastgele Etki Modeli				
Denekler arası (DA)	n-1	KOD _A	$k\sigma_d^2 + \sigma_h^2$	$\frac{\sigma_d^2}{\sigma_d^2 + \sigma_h^2}$
Denekler içi (Di)	n (k-1)	KOD _i	σ_h^2	
İki Yönlü Rastgele Etki Modeli				
Denekler arası (DA)	n-1	KOD _A	$k\sigma_d^2 + \sigma_h^2$	$\frac{\sigma_d^2}{\sigma_d^2 + \sigma_g^2 + \sigma_h^2}$
Denekler içi (Di)	n (k-1)	KOD _i	$\sigma_g^2 + \sigma_h^2$	
Gözlemciler arası (GA)	k-1	KOG _A	$n\sigma_g^2 + \sigma_h^2$	
Hata	(n-1) (k-1)	KOH _A	σ_h^2	
İki Yönlü Karma Etki Modeli				
Denekler arası (DA)	n-1	KOD _A	$k\sigma_d^2 + \sigma_h^2$	$\frac{\sigma_d^2}{\sigma_d^2 + \theta_g^2 + \sigma_h^2}$
Denekler içi (Di)	n (k-1)	KOD _i	$\theta_g^2 + \sigma_h^2$	
Gözlemciler arası (GA)	k-1	KOG _A	$n\theta_g^2 + \sigma_h^2$	
Hata	(n-1) (k-1)	KOH _A	σ_h^2	

TABLO 3: Tutarlılık ve mutlak uyum tanımına göre farklı veri setleri için elde edilen SKK tahminleri.

<p>Örnek 1:</p> <p>İkinci gözlemci her zaman birinci gözlemciden 4 birim daha yüksek ölçmektedir.</p> <p>(1,5); (2,6); (3,7); (4,8); (5,9).</p>	
<p>Örnek 2:</p> <p>İkinci gözlemci her zaman birinci gözlemciden 1.5 kat daha yüksek ölçmektedir.</p> <p>(1,1.5); (2,3); (3,4.5); (4,6); (5,7.5).</p>	
<p>Örnek 3:</p> <p>İkinci gözlemci her zaman birinci gözlemci ile aynı ölçmektedir.</p> <p>(1,1); (2,2); (3,3); (4,4); (5,5).</p>	

KULLANILAN ÖLÇÜMLERİN ELDE EDİLME BİÇİMİNE GÖRE SKK TAHMİNLERİ

SKK hesaplaması, gözlem değerlerinin tek bir ölçüm sonucuna mı yoksa birden çok ölçümün toplamı ya da ortalamasına göre mi hesaplandığına bağlı olarak değişir. Örneğin iki diş hekiminin bir dolgu maddesinin yayılımını ölçmek için yaptığı bir çalışmada dolgu maddesi ile diş arasında kalan boşluk miktarlarını ölçtüklerini varsayalım. Gözlemciler arası uyum her bir diştten alınan tek bir kesite göre inceleniyorsa “tek (single)”, her bir diştten 3 kesit (apikal, middle, korneal) alınarak dişi temsilen üç kesitin ortalama değerleri üzerinden uyum

inceleniyorsa “ortalama (average)” seçeneği kullanılmalıdır. “Ortalama” seçeneğinden elde edilen SKK değerleri genellikle “tek (single)” seçeneğinden elde edilen SKK değerlerinden daha yüksektir. Çünkü birden fazla ölçümden elde edilen sonuçlar genellikle daha güvenilir sonuçlar verir.⁷

“Tek (single)” ve “ortalama (average)” ayrımı, yukarıda bahsedilen her üç model için de mevcuttur. Bu modellerde ortalama seçeneğine bağlı olarak SKK tahmini hesaplanırken, denekler içi varyans terimi kullanılan ölçüm sayısına bağlı olarak düzeltilir.

Çalışmanın tasarımına, amacına ve kullanılan ölçümlerin elde edilme biçimine bağlı olarak he-

TABLO 4: Çalışmanın tasarımına, amacına ve ölçümlerin elde edilme biçimine bağlı olarak hesaplanabilecek farklı SKK türleri.

SKK Türleri	Tanım
SKK(1,1)	<ul style="list-style-type: none"> Her bir deneğin her bir ölçümü rastgele seçilmiş gözlemciler tarafından yapılır. SKK tek bir ölçüm kullanılarak hesaplanır.
SKK(1,k)	<ul style="list-style-type: none"> Her bir deneğin her bir ölçümü rastgele seçilmiş gözlemciler tarafından yapılır. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.
SKK(2,1) _C	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemciler rastgele seçilmiştir. SKK tek bir ölçüm kullanılarak hesaplanır.
SKK(2,1) _A	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemciler rastgele seçilmiştir. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.
SKK (2,k) _C	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemciler rastgele seçilmiştir. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.
SKK (2,k) _A	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemciler rastgele seçilmiştir. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.
SKK (3,1) _C	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemci etkisi sabittir. SKK, tek bir ölçüm kullanılarak hesaplanır.
SKK (3,1) _A	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemci etkisi sabittir. SKK, tek bir ölçüm kullanılarak hesaplanır.
SKK (3,k) _C	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemci etkisi sabittir. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.
SKK (3,k) _A	<ul style="list-style-type: none"> Her bir gözlemci her bir denekten ölçüm alır. Gözlemci etkisi sabittir. SKK, "k" ölçümün toplamı ya da ortalaması kullanılarak elde edilir.

saplanabilecek olası tüm SKK tahminleri Tablo 4'te özetlenmiştir. Bu tabloda SKK değerlerine ilişkin ilk indis model türlerini (1: Tek yönlü rastgele etki modeli, 2: İki yönlü rastgele etki modeli, 3: İki yönlü karma etki modeli), ikinci indis kullanılan ölçümlerin elde edilme biçimini (1: tek ölçüm, k: toplam/ortalama ölçüm) göstermektedir. SKK (*,*)_C, tutarlılık tanımına göre, SKK (*,*)_A ise mutlak uyum tanımına göre hesaplanan SKK tahminini ifade etmektedir.

Ölçüm değerlerinin sürekli olduğu durumda, gözlemci içi ve gözlemciler arası uyumun değerlendirilmesinde en yaygın kullanılan güvenilirlik ölçü-

sü SKK'dır. Temelleri ilk olarak Fleiss ve Shrout tarafından atılan SKK'nın farklı deney düzenlerinde kullanılmak üzere geliştirilmiş çok sayıda formu bulunmakta ve bunlar SPSS, SAS gibi paket programlar yardımıyla da kolaylıkla hesaplanabilmektedir. Ancak sağlık alanında yapılan çalışmalar incelendiğinde, veri setine en uygun SKK modelinin doğru olarak belirlenemediği gözlemlendiğinden, çalışmamızda SKK tahminlerinin değişkenlik göstermesinde etkili olan faktörler detaylı olarak incelenmiştir. Bu çalışmanın sağlık alanında güvenilirliğin değerlendirilmesi konusunda araştırma yapan kişiler için bir kılavuz niteliği taşıması hedeflenmektedir.

KAYNAKLAR

- Alpar R. [Performance tests of validity and reliability]. Spor Bilimlerinde Uygulamalı İstatistik. 2nd ed. İstanbul: Nobel Yayın Dağıtım; 2001. p.261.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420-8.
- Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989;45(1):255-68.
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1(1):30-46.
- Shoukri MM, Pause CA. Statistical analysis of measurements reliability. Statistical Methods for Health Sciences. 2nd ed. Florida: CRC Press LLC; 2000. p.19-41.
- Bland JM, Altman DG. Measurement error and correlation coefficients. BMJ 1996;313(7048):41-2.
- Nichols DP. Choosing an intraclass correlation coefficient. SPSS Keywords 1998;67:1-2.
- Haber M, Barnhart HX. Coefficients of agreement for fixed observers. Stat Methods Med Res 2006;15(3):255-71.