



**ALTIN STANDART TESTİN OLMADIĞI DURUMDA TANI TESTLERİNİN
PERFORMANSLARININ DEĞERLENDİRİLMESİ**

**Assessing the Performance of Diagnostic Tests When a Gold
Standard Does Not Exist**

¹ Semra Erdoğan

¹ Gülhan Örekici Temel

ÖZET

Klinikte belirli bir hastalığın tanısının konulmasında, altın standart testin olmadığı ama iki veya daha fazla kesin olmayan referans testin var olduğu durumlarla karşılaşmaktadır. Yeni testin kesin olmayan referans testlere göre performanslarını değerlendirmek için literatürde Discrepant Resolution (DR), Composite Referans Standart (CRS) ve Latent Class model (LCA) olmak üzere üç temel yaklaşım bulunmaktadır. Bu çalışma ile üç temel yaklaşımın değerlendirilmesi amaçlanmıştır. LCA, DR ve CRS yaklaşımından çok daha karmaşık bir yapıya sahiptir. LCA, çözücü test ya da diğer bir kesin olmayan referans testin seçiciliği hakkında herhangi bir varsayım gerektirmeksizin CRS ve DR yöntemlerinden daha avantajlıdır. Sonuç olarak, yeni tanı testinin performansını değerlendirmek ve kesin olmayan referans testler sonucunda ortaya çıkan yanlışlığı azaltmak adına geliştirilen yöntemlerden LCA, diğer yöntemlere nazaran daha güvenilir sonuçlar vermektedir.

Anahtar kelimeler: Tanı testi, Composite referans standart, Gizli sınıf analizi.

ABSTRACT

Situations are encountered that gold standard doesn't exist and two or more imprecise reference tests are available in the diagnosis of a specific disease clinically. There are three basic approaches including Discrepant Resolution (DR), Composite Reference Standard (CRS) and Latent Class model (LCA) in literature for assessing the performance of diagnostic tests. This study is aimed to evaluate the three basic approaches. LCA has a more complex structure than DR and CRS approach. LCA is more advantageous than CRS and DR methods without any assumptions about resolver test or other imperfect reference tests specify. As a result, LCA methods which are developed for assessing the performance of new diagnostic test and reducing the bias because of imperfect reference tests give more reliable results than other methods.

Key words: Diagnostic test, Composite reference standard, Latent class analysis.

GİRİŞ

Tanı testleri, hasta ve sağlıklı bireylerin oluşturduğu popülasyondan rasgele alınan bireylerin hasta ya da sağlıklı olup olmadıklarının doğru olarak saptanması amacıyla kullanılan yöntemler dizisidir. Herhangi bir hastalık ile ilgili olarak tanı koymada kullanılan birden fazla test olabilir, ancak bu testlerden sadece birisi ile bireylere kesin hasta ya da kesin sağlıklı tanıları konulmaktadır ki bu test altın standart test (gold standart, referans gold standart test) olarak adlandırılmaktadır. Hasta ve sağlıklı olmak üzere iki sonuçlu tanı testinin olduğu durumlarda ayırt etme gücü “duyarlılık”, “seçicilik”, “pozitif tahmini değer”, “negatif tahmini değer” gibi ölçütlerle, sıralı ya da sürekli sonuçlu tanı testinin olduğu durumlarda ise ayırt etme gücü “eğri altında kalan alan” yardımı ile değerlendirilmektedir (1).

Sağlık hizmetlerinde tanı testleri önemli bir rol oynamaktadır. Özellikle viral ve bakteriyel enfeksiyonları ayırt edebilmek için sürekli olarak yeni bir tanı testinin geliştirilmesi gerçekleştirilmektedir. Yeni geliştirilen bir tanı testinin performansı da enfeksiyon durumunu kesin olarak değerlendirebilen bir altın standart testin sonuçları ile karşılaştırılmaktadır. Ancak, bazı hastalıkların test edilmesinde herhangi bir altın standart testin olmaması, altın standart test olduğu halde uygulamasının zor ve masraflı, teknik olarak uygun olmaması ya da tedaviye başlamak için altın standart testin sonucunun beklenmesinin hastanın hayatını riske atılmasına sebep olması gibi durumlardan ötürü altın standart testin kullanılmadığı durumlarla karşılaşılabilir. Böyle durumlarda tanı testinin doğruluk ölçütleri, göreceli olarak diğerlerine göre en doğru sonuçları veren fakat sonuçları kesin doğru olmayan bir teste göre hesaplanır ki bu tür testler “kesin olmayan referans testler (imperfect gold standard)” olarak adlandırılır (2-

¹ Mersin Üniversitesi Tıp
Fakültesi, Biyoistatistik ve Tıbbi
Bilişim Anabilim Dalı, Mersin

Submitted/Başvuru tarihi:
17.05.2012
Accepted/Kabul tarihi:
05.03.2015
Registration/Kayıt no:
12.05.266

**Corresponding Address /
Yazışma Adresi:**

Semra Erdoğan

Mersin Üniversitesi, Tıp Fakültesi
Dekanlık Binası Biyoistatistik ve
Tıbbi Bilişim AD, Çiftlikköy
Kampüsü, Mersin.

Tel: 0324 3610001- 1029
Fax: 0324 3412400

E-posta:
semraerdogann@gmail.com

© 2012 Düzce Medical Journal
e-ISSN 1307- 671X
www.tipdergi.duzce.edu.tr
duzcetipdergisi@duzce.edu.tr

4). Metodolojide duyarlılığı ve seçiciliği bilinen ya da duyarlılık ve seçiciliği başka çalışmalardan tahmin edilebilen kesin olmayan bir referans test kullanılarak tanı testinin doğruluk ölçütlerinin tahmin edilmesi önerilmektedir (5,6). Ancak, pek çok araştırmacı, kesin olmayan referans test kullanılarak tanı testinin doğruluk ölçütlerinin tahmin edilmesinde yanlışlığın varlığını tartışmaktadırlar (7-9). Ayrıca, yeni tanı testi ile kesin olmayan referans test koşullu bağımsız olduğunda testin duyarlılığı ve seçiciliği gerçekte olduğundan daha düşük, koşullu bağımlı olduğunda ise duyarlılık ve seçiciliğin gerçek değerinden daha yüksek tahmin edildiği ifade edilmektedir (4).

Klinikte, belirli bir hastalığın tanısının konulmasında, altın standart testin olmadığı ama iki veya daha fazla kesin olmayan referans testlerin bir arada kullanıldığı durumlarla karşılaşmaktadır. Kesin olmayan referans testin duyarlılık ve seçiciliğinin olmadığı veya başka çalışmalardan tahmin edilemediği ve bundan dolayı ortaya çıkan yanlışlığı azaltmak amacıyla literatürde karşılaşılan pek çok çalışma bulunmaktadır. Bu çalışmalardan yola çıkarak, Discrepant Resolution (DR), Composite Referans Standart (CRS) ve Latent Class model (LCA) olmak üzere literatürde yer alan üç temel yaklaşım bulunmaktadır (3,10-12).

DR yöntemi altın standart testin olmadığı durumlarda yeni tanı testinin doğruluğunu tahmin edebilmek için kullanılan bir yöntemdir ancak bu yöntemde pek çok farklı kesin olmayan referans testler mevcuttur. Tek bir kesin olmayan referans testin kullanıldığı ya da DR yönteminin kullanıldığı durumlara alternatif olarak birden fazla kesin olmayan testin var olabileceği temeline dayanan CRS yöntemi önerilmiştir. DR ve CRS yöntemine alternatif olarak, klinik tanı testlerinin doğruluk ölçütlerinin değerlendirilmesinde kullanılan, ayrıca kesin olmayan referans testin duyarlılık ve seçicilik değerlerindeki yanlışlıkları azaltan LCA yaklaşımı geliştirilmiştir (12).

Bu çalışma ile altın standart testin olmadığı ama kesin olmayan referans testin kullanıldığı durumlarda, kesin olmayan referans testin kullanımına bağlı olarak ortaya çıkan yanlışlığı ortadan kaldırmak için, tanı testlerinin performanslarını değerlendirmek üzere geliştirilmiş olan ve yukarıda sözü edilen üç temel yaklaşımın değerlendirilmesi amaçlanmıştır.

Discrepant Resolution (DR)

DR yönteminin amacı, altın standart olmayan bir referans test ile yeni tanı testinin karşılaştırılması ve bu yeni tanı testine ait doğruluk ölçütlerinin elde edilmesidir. Bu yöntemle, özellikle, yeni tanı testinin ve kesin olmayan referans testin birbirinden farklı olduğu (uyuşmadığı) durumlarda, çözücü (resolver) test olarak adlandırılan üçüncü bir test kullanılarak referans testin kaçırdığı gerçek pozitif durumlar yakalanmaya çalışılmaktadır (11,12).

DR yöntemi 1984 yılından beri kullanılmaktadır. DR algoritması iki aşamadan oluşmaktadır (Tablo 1). İlk aşamada ilk kullanılan kesin olmayan referans test (Test 1) altın standart test (referans) olarak düşünülerek, yeni test sonucu referans teste göre test edilir (Tablo 1, 1. Aşama).

Tablo 1: Discrepant Resolution yöntemi ile elde edilen iki aşamalı birleştirilmiş çapraz tablo.

		1. Aşama		2. Aşama	
		Test 1		(DR)	
		+	-	+	-
Yeni Test	+	n_{++}	n_{+-}	$n_{++} + n_{+-}$	n_{+--}
	-	n_{-+}	n_{--}	n_{-++}	$n_{--} + n_{-+-}$

Çapraz tabloda her bir hücre test sonuçlarının bir kombinasyonudur. İlk alt simge yeni test sonucunu ikinci alt simge ise test 1 sonucunu göstermektedir. n_{++} , her iki test sonucunun da pozitif olduğu duruma ait birey sayısını gösterirken, yeni test sonucu pozitif, test 1 sonucu negatif olan duruma ait birey sayısını göstermektedir.

Test 1 referans test olarak alındığında yeni testin doğruluk ölçütleri olan duyarlılık (sensitivity), seçicilik (specificity), pozitif tahmin değeri (PPV), negatif tahmin değeri (NPV) ve prevalans (Pr) değerleri Tablo 1'in 1. Aşamadaki verilerinden yararlanılarak sırasıyla Eşitlik 1-5'deki gibi hesaplanmaktadır (12).

$$Sensitivity = \frac{n_{++}}{n_{++} + n_{+-}} \quad (1)$$

$$Specificity = \frac{n_{--}}{n_{--} + n_{-+}} \quad (2)$$

$$PPV = \frac{n_{++}}{n_{++} + n_{-+}} \quad (3)$$

$$NPV = \frac{n_{--}}{n_{--} + n_{+-}} \quad (4)$$

$$Pr = \frac{n_{++} + n_{-+}}{n} \quad (5)$$

İkinci aşamada, test 1 ile yeni testin uyuşmadığı (yeni test pozitif kararı verirken test 1'in negatif kararı verdiği ya da yeni test negatif kararı verirken test 1 sonucunun pozitif kararı verdiği durumlar) bireylere kesin olmayan üçüncü bir referans test yeni çözücü bir test uygulanır. Bu test sonucuna göre bireyler yeniden hasta ya da sağlam olarak sınıflandırılır ve yeni bir çapraz tablo elde edilir (Tablo 1, 2. Aşama).

Çapraz tabloda üçüncü alt simgeler ise çözücü test ile yeniden test edildiğinde elde edilen birey sayılarını göstermektedir. Yeni test sonucu pozitif, test 1 sonucu negatif olan (n_{+-}) bireylere çözücü test uygulandığında sonuç pozitif ise ile gösterilmekte ve hasta olarak sınıflandırılmaktadır, ancak çözücü test sonucu negatif olan bireyler ile gösterilmekte ve sağlam olarak sınıflandırılmaktadır. Yeni test sonucu negatif, test 1 sonucu pozitif olan bireyler içinde tekrarlanır. DR algoritması kullanıldığında ise duyarlılık (sensitivity), seçicilik (specificity), pozitif tahmin değeri (PPV), negatif tahmin değeri (NPV) ve prevalans değeri sırasıyla Eşitlik 6-10'daki gibi hesaplanmaktadır (12).

$$Sensitivity_{DR} = \frac{n_{++} + n_{-+-}}{n_{++} + n_{-+-} + n_{-++}} \quad (6)$$

$$Specificity_{DR} = \frac{n_{--} + n_{-+-}}{n_{--} + n_{-+-} + n_{-++}} \quad (7)$$

$$PPV_{DR} = \frac{n_{++} + n_{-+-}}{n_{++} + n_{-+-} + n_{-++}} \quad (8)$$

$$NPV_{DR} = \frac{n_{--} + n_{-+-}}{n_{--} + n_{-+-} + n_{-++}} \quad (9)$$

$$Pr_{DR} = \frac{n_{++} + n_{-+-} + n_{-++}}{n} \quad (10)$$

DR yöntemi ile ilgili önemli bir sorun, elde edilen duyarlılık ve seçicilik tahminlerinin yorumlanmasının belirsiz olması ile ilgilidir. Doğruluğun ölçülebilirliği ile ilgili olarak, yeni testin duyarlılık ve seçicilik formülasyonlarında (Eşitlik 6 ve 7) görüldüğü gibi doğal olarak yeni testin sonuçları ile ilişkilidir. Yeni teste ait doğruluk değerlendirildiğinde, yeni testin bağımsız olduğu bir standartla karşılaştırılması zorunludur. DR yönteminde bu prensip ihmal edilmektedir (11,12). Yani, yeni testin doğruluk ölçütleri, kesin olmayan referans testin ve yeni testin uyuşmadığı hücreler üzerinden hesaplandığı ve uyumlu olan hücreler çözücü test ile incelenmediği için kesin olmayan referans teste bağlı olarak ortaya çıkan yanlışlık önlenememektedir. DR yöntemi, tanı

testlerinin performanslarını değerlendirmek için çok yaygın olarak kullanılmasına rağmen, Miller (13), Hadgu (14), Lipman ve Astles (15), duyarlılık ve seçicilik tahminlerindeki yanlışlıklardan ötürü yeni teste ait duyarlılık ve seçicilik değerlerinin gerçekte olduğundan çok daha yüksek olduğunu göstermişlerdir (10,12-16).

DR yöntemine göre referans alındığında yeni teste ait duyarlılık, seçicilik, PPV ve NPV değerleri, kesin olmayan referans teste (Test 1) göre yeni testin hesaplanan duyarlılık, seçicilik, PPV ve NPV değerleri eşit veya daha büyük olacaktır. Bu durum yalnızca, Tablo 1'in 2. aşamasında yani DR yöntemindeki köşegenlerde yer alan bireylerin sayısı arttığı zaman değişebilir (11,12).

Composite Referans Yöntemi (CRS)

Altın standardın olmadığı, birden fazla referans testin var olduğu tüm durumlarda kullanılabilir bir yöntemdir. Bu yöntem DR yönteminde yer alan temel probleme karşılık alternatif olarak geliştirilmiştir. DR yönteminde olduğu gibi bu yöntem de iki aşamada gerçekleştirilmektedir. DR yönteminde olduğu gibi, ilk aşamada yeni test, test 1 ile test edilir (Tablo 2, 1. Aşama). İkinci aşamada test 2'nin pozitif ve negatif kararlarına göre çapraz tablo CRS yöntemi için yeniden düzenlenir (Tablo 2, 2. Aşama). İlk aşamada tüm test 1'e ait pozitif durumlar ikinci aşamada da CRS yönteminde pozitif olarak tanımlanır iken yalnızca ilk aşamada Test 1'e ait negatif durumlar ikinci aşamada çözücü test ile test edilmektedirler. Yani, DR yöntemi ile referans testin ve yeni testin uyumadığı hücreler üzerinden yeni testin doğruluk ölçütleri hesaplanırken, CRS yönteminde, kesin olmayan referans testin negatif olan tüm örneklemi üzerinden yeni testin doğruluk ölçütleri hesaplanmaktadır. Kesin olmayan referans test (test 1) veya yeni test pozitif olduğu durumlarda CRS pozitif, Test 1 ve yeni test negatif olduğu durumlarda ise CRS negatif olarak tanımlanmaktadır. n+-, yeni test sonucu pozitif, birinci test sonucu negatif ve çözücü test sonucunda da pozitif olduğu durumdaki birey sayısını göstermektedir (10-12,17-18). Bu durumda yeni testin diagnostik doğruluk istatistikleri CRS yöntemine göre Eşitlik 11-15'deki gibi hesaplanmaktadır (12).

$$Sensitivity_{CRS} = \frac{n_{++} + n_{+-}}{n_{++} + n_{+-} + n_{+0} + n_{+1}} \quad (11)$$

$$Specificity_{CRS} = \frac{n_{--}}{n_{--} + n_{-0} + n_{-1}} \quad (12)$$

$$PPV_{CRS} = \frac{n_{++} + n_{+-}}{n_{++} + n_{+-} + n_{+0} + n_{+1}} \quad (13)$$

$$NPV_{CRS} = \frac{n_{--}}{n_{--} + n_{-0} + n_{-1}} \quad (14)$$

$$Pr_{CRS} = \frac{n_{++} + n_{+-} + n_{-0} + n_{-1}}{n} \quad (15)$$

Tablo 2: Composite referans yöntemi ile elde edilen iki aşamalı birleştirilmiş çapraz tablo.

		Test 1		Test 2		
				CRS		
		+	-	+	-	-
Yeni Test	+	n ₊₊	n ₊₊ + n ₊₀	n ₊₊	n ₊₀	n ₊₁
	-	n ₋₀	n ₋₀ + n ₋₁	n ₋₀	n ₋₀	n ₋₁

CRS, ikiden fazla kesin olmayan referans test için de genişletilebilir. Kesin olmayan referans test sayısındaki artışa bağlı olarak ilgilenilen durumun varlığı ya da yokluğunu belirlemek üzere geliştirilen karar kuralı değişir. Eğer hastalığın tanısını koymada altı tane kesin olmayan referans testten yararlanılıyorsa en az bir test sonucunun pozitif çıkması durumunda sonuç pozitif olarak yorumlanabilir ya da iki veya daha fazla testin sonucu pozitif olduğunda ilgilenilen durumun varlığı konusunda karar verilebilir (19).

CRS yönteminde, yalnızca kesin olmayan referans testin pozitif durumundaki denek sayılarını değil aynı zamanda çözücü testin pozitif durumundaki denek sayılarını da etkilenmektedir. Bu nedenle, CRS yöntemine göre elde edilen prevalans değerinin kesin olmayan referans testin (test 1) altın standart olarak kabul edildiği durumdaki prevalans değerine eşit ya da daha yüksek bulunması sürpriz olmamaktadır. Ayrıca, CRS yöntemine göre yeni testin PPV değeri Test 1 referans test olarak kabul edildiğinde yeni testin PPV değerine eşit ya da daha büyük olacaktır. Bunun yanı sıra CRS yöntemine göre yeni testin NPV değeri de Test 1 referans test olarak kabul edildiğinde yeni testin NPV değerine eşit ya da daha küçük olacaktır (12). Eğer Test 1'in referans test kabul edildiği durumda yeni testin duyarlılık değeri $\frac{n_{++}}{n_{++} + n_{+-}}$ değerine eşit ve büyük ise Test 1'in referans test kabul edildiği durumda yeni testin duyarlılık değeri CRS yöntemine göre yeni testin duyarlılık değerine eşit ve daha büyük olacaktır (12).

Test 1'in referans test kabul edildiği durumda yeni testin seçiciliği $\frac{n_{--}}{n_{--} + n_{-0} + n_{-1}}$ değerine eşit ve büyük ise Test 1'in referans test kabul edildiği durumda yeni testin seçicilik değeri CRS yöntemine göre yeni testin seçicilik değerine eşit ve daha küçük olacaktır (12).

DR ve CRS yöntemleri karşılaştırıldığında, CRS yöntemi kullanılarak hesaplanan prevalans değeri DR yöntemi kullanılarak hesaplanan prevalans değerine eşit ve daha büyüktür. Diğer taraftan, CRS yöntemine göre elde edilen duyarlılık, seçicilik ve NPV değerleri DR yönteminden elde edilen duyarlılık, seçicilik ve NPV değerlerine eşit ve daha küçüktür. İlginçtir ki, her iki yöntemde ait PPV formülasyonları (Eşitlik 8 ve 13) birbirine eşittir ve her iki yöntemden elde edilen PPV sonuçları da eşittir (12).

Latent Class Analizi (Gizli Sınıf Analizi, LCA)

Duyarlılığı ve seçiciliği yüksek olan bir altın standardın yokluğunda bir tanı testinin doğruluğunu test edebilmek için kullanılan, ayrıca kesin olmayan referans testin duyarlılık ve seçicilik değerlerindeki yanlışlıkları azaltan CRS yöntemine alternatif ikinci bir geçerli yaklaşımdır. DR ve CRS yöntemlerinin aksine minimum üç kesin olmayan referans test gereklidir (10-12).

LCA'da hastalığın tanısına yönelik olarak kullanılan mevcut bütün tanı testleri, kesin olmayan referans test olarak kabul edilmekte ve bu gözlenen tanı testlerinin sonuçlarına dayanılarak oluşturulan matematiksel model yardımıyla gözlenemeyen gerçek hastalık durumu kestirilmektedir (12).

Bir altın standardın olmadığı durumda yeni geliştirilen tanı testinin doğruluğu, ya maksimum olabilirlik tahminleri (MLE) ile ya da Bayesian yaklaşımı (BA) ile test edilmektedir. Maksimum olabilirlik tahminlerinin kullanımı ilk olarak 1980 yılında Hui ve Walter tarafından gerçekleştirilmiştir. Hui ve Walter'a göre, LCA'e ait önemli yaklaşımlar bulunmaktadır ve bu yaklaşımlar Hui ve Walter paradigması olarak bilinmektedir. Bu yaklaşıma göre, çalışılan populasyonun, farklı prevalanslara sahip iki (veya daha fazla) alt populasyondan oluştuğu varsayılmaktadır. Her birey yalnızca tek bir alt populasyona ait olmakta ve gözlenemeyen her bir sınıfta yer alan değişkenler istatistiksel olarak birbirinden bağımsız olmaktadır (10,20,21).

LCA'da, bu iki alt populasyonda, gözlenemeyen gizli iki kategorili

değişken D ile gösterilmekte ve genellikle 0 (D=0, hastalık yok) ve 1 (D=1, hastalık var) şeklinde kategorize edilmektedir. $i=1, \dots, n$ 'e kadar her bir hastanın J tane iki kategorisi olan tanı testi ile sınıflandırılması sonuçları $Y_i (Y_{i1}, \dots, Y_{iJ})$ ile gösterilmektedir. Koşullu bağımsızlık varsayımı altında \emptyset parametresi ile Y_i 'nin birleşik dağılımına ilişkin istatistiksel model Eşitlik 16'daki gibi verilmektedir (22).

$$P_{\theta}(Y_{i1}, Y_{i2}, \dots, Y_{iJ}) = \sum_{I=0}^1 P_{\theta}(Y_{i1}, Y_{i2}, \dots, Y_{iJ} / D_i = I) P(d_i = I) \quad (16)$$

j. teste ait doğru ve yanlış pozitif oranları sırasıyla Eşitlik 17 ve 18'de verilmektedir (22).

$$\phi_j = P(Y_{ij} = 1 / D_i = 1) \quad (17)$$

$$\psi_j = P(Y_{ij} = 1 / D_i = 0) \quad (18)$$

Eğer model anlamlı ise $\theta = \{\phi_j, \psi_j\}$, $j=1, \dots, J$ ve $\rho = P(D_i = 1)$ hastalık prevalansı Eşitlik 19'deki olabilirlik fonksiyonu ile maksimize edilerek tahmin edilmektedir (22).

$$L(\theta, \rho) = \prod_{i=1}^n \left\{ \rho P_{\theta}(Y_{i1}, \dots, Y_{iJ} / D_i = 1) + (1 - \rho) P_{\theta}(Y_{i1}, \dots, Y_{iJ} / D_i = 0) \right\} \quad (19)$$

$P_{\theta}(Y_{i1}, Y_{i2}, \dots, Y_{iJ})$ için en basit ve popüler model koşullu bağımsızlık varsayımı altında Eşitlik 20'deki gibi verilmektedir (22).

$$L(\theta, \rho) = \prod_{i=1}^n \left\{ \rho \prod_{j=1}^J P(Y_{ij} / D_i = 1) + (1 - \rho) \prod_{j=1}^J P(Y_{ij} / D_i = 0) \right\} \quad (20)$$

LCA'nın tanımlı olabilmesi için yeterli koşul en çok olabilirlik kestirimleri için elde edilen kovaryans matrisinin pozitif tanımlı olması, gerekli koşul ise serbestlik derecesinin negatif olmamasıdır. LCA'da J tane iki kategorisi olan tanı testi için 2J tane pozitif ve negatif sınıflandırma yapılmaktadır. Bağımsız parametre sayısı (2J+1) ve serbestlik derecesi (2J - 1) şeklinde hesaplanmaktadır. Üç kesin olmayan tanı testi kullanılıyorsa (J=3), bağımsız parametre sayısı 7 (1 tane prevalans, 3 tane duyarlılık, 3 tane seçicilik) ve serbestlik derecesi 7 olacaktır (22).

LCA, uygulamada çoğu zaman ihmal edilen doğrusal ilişki, normal dağılım ve homojenlik gibi geleneksel modelleme varsayımlarına dayanmamaktadır. Bu nedenle model varsayımlarının sağlanamamasından kaynaklanan sapmalara daha az duyarlıdır (23).

LCA, birkaç alanda eleştirilmektedir. Yapılan çalışmalarda, koşullu bağımsızlık varsayımı göz ardı edildiğinde tanı testine ait doğruluk ölçütlerinin yanlış olarak tahmin edildiği ve modelde koşullu bağımsızlık varsayımının üç kesin olmayan referans test kullanımında test edilemediği gözlenmiştir (12,19).

LCA, DR ve CRS yaklaşımından çok daha karmaşık bir yapıya sahiptir. Çünkü LCA, gözlenemeyen veriye modelleyebilen uygun bir istatistik model içermektedir. LCA, çözücü test ya da diğer bir kesin olmayan referans testin seçiciliği hakkında herhangi bir varsayım gerektirmeksizin, CRS ve DR yöntemlerinden daha avantajlıdır (11).

Sonuç olarak, altın standart testin olmadığı ama kesin olmayan referans testin kullanıldığı durumlarda, yeni tanı testinin performansını değerlendirmek ve kesin olmayan referans testler sonucunda ortaya çıkan yanlışlığı azaltmak adına geliştirilen yöntemlerden LCA, diğer yöntemlere nazaran daha güvenilir sonuçlar vermektedir.

KAYNAKLAR

1. Reitsma JB, Rutjes AWS, Khan KS, Coomarasamy A, Bossuyt PM. A Review of solution for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol*. 62: 797-806, 2009. DOI: 10.1016/j.jclinepi.2009.02.005.
2. Genç Y. Tanı testi çalışmalarında metodolojik standartların kullanılması. Ankara Üniversitesi Tıp Fakültesi Mecmuası. 56

(4): 259-64, 2003.

3. Albert PS. Estimating diagnostic accuracy of multiple binary tests with an imperfect reference standard. *Stat Med*. 28: 780-97, 2009. DOI: 10.102/sim.3514.
4. Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. Wiley Series in Probability and Statistics. New York: John Wiley & Sons, Inc. pp:71-77, 2002.
5. Staquet M, Rozenzweig M, Lee YJ, Muggia FM. Methodology for the assessment of new dichotomous diagnostic tests. *Journal of Chronic Diseases*. 34:599-610, 1981. [http://dx.doi.org/10.1016/0021-9681\(81\)90059-X](http://dx.doi.org/10.1016/0021-9681(81)90059-X).
6. Baker SG. Evaluating a new test using a reference test with estimated sensitivity and specificity. *Communications in Statistics: Theory and Methods*. 20:2739-52, 1991.
7. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med*. 140(3):189-202, 2004.
8. Valenstein PN. Evaluating diagnostic tests with imperfect standards. *Am J Clin Pathol*. 93(2):252-8, 1990.
9. Begg CB. Biases in the assessment of diagnostic tests. *Stat Med*. 6:411-423, 1987.
10. Nerette P, Stryhn H, Dohoo I, Hammel L. Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests. *Prev Vet Med*. 85: 207-25, 2008. DOI: 10.1016/j.prevetmed.2008.01.011.
11. Baughman AL, Bisgard KM, Cortese MM, Thompson WW, Sanden GN, Strebel PM. Utility of composite reference standards and latent class analysis in evaluating the clinical accuracy of diagnostic tests for pertussis. *Clin Vaccine Immunol*. 15 (1): 106-14, 2008. DOI:10.1128/0014-8177.00223-07.
12. Alonzo TA, Pepe MS. Using a combination of reference tests to assess the accuracy of a new diagnostic test. *Stat Med*. 18: 2987-3003, 1999.
13. Miller W. Editorial response: can we do better than discrepant analysis for new diagnostic test evaluation? *Clin Infect Dis*. 27(5): 1186-93, 1998.
14. Hadgu A. The discrepancy in discrepant analysis. *Lancet*. 348 (9027): 592-93, 1996. DOI: [http://dx.doi.org/10.1016/S0140-6736\(96\)05122-7](http://dx.doi.org/10.1016/S0140-6736(96)05122-7).
15. Lipman HB, Astles JR. Quantifying the bias associated with use of discrepant analysis. *Clin Chem*. 44(1): 108-15, 1998.
16. Hawkins MD, Garrett AJ, Stephenson B. Some issues in resolution of diagnostic tests using an imperfect gold standard. *Stat Med*. 20: 1987-2001, 2001.
17. Jang D, Sellors JW, Mahony JB, Pickard L, Chernesky MA. Effects of broadening the gold standard on the performance of a chemiluminometric immunoassay to detect Chlamydia trachomatis antigens in centrifuged first void urine and urethral swab samples from men. *Sex Transm Dis*. 19(6): 315-9, 1992.
18. Hess AS, Shardell M, Johnson JK, Thom KA, Strassle P, Netzer G, et al. Methods and recommendations for evaluating and reporting a new diagnostic test. *Eur J Clin Microbiol Infect Dis*. Published online: 29 March 2012. DOI: 10.1007/s100096-012-1602-1.
19. Rutjes AWS, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Methodological research in diagnostic tests: Evaluation of diagnostic tests when there is no gold standard-A review of methods. NHS Research Methodology Programme 2006, UK (grant: RM04/JH21: No Gold Standard).
20. Toft N, Jorgensen E, Hojsgaard S. Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of

sensitivity and specificity in the absence of a gold standard. *Prev Vet Med.* 68: 19-33, 2005. <http://dx.doi.org/10.1016/j.prevetmed.2005.01.006>.

21. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics.* 36:167–171, 1980.
22. Pepe MS, Janes H. Insights into latent class analysis of diagnostic test performance. *Biostatistics.* 8(2): 474-84, 2007. DOI: 10.1093/biostatistics/kxl038.
23. Vermunt JK. Latent Class and finite mixture models for multilevel data sets. *Stat Methods Med Res.* 1-19, 2007. DOI:10.1177/0962280207081238.