



T.C.

**SÜLEYMAN DEMİREL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM ANABİLİM DALI**

**KLİNİK VAKA TANISINDA
ROC ANALİZİ TEKNİĞİ**

Ebru DUMLUPINAR

YÜKSEK LİSANS TEZİ

DANIŞMAN

Dr. Öğr. Üyesi Osman GÜRDAL

Tez. No: 171

ISPARTA-2019

KABUL ve ONAY SAYFASI

Sağlık Bilimleri Enstitü Müdürlüğüne;

Süleyman Demirel Üniversitesi Sağlık Bilimleri Enstitüsü **Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı Yüksek Lisans Programı** çerçevesinde yürütülmüş olan bu çalışma, aşağıdaki jüri tarafından **Yüksek Lisans Tezi** olarak kabul edilmiştir.

Tez Savunma Tarihi: 28/05/2019

Tez Danışman : Dr. Öğr. Üyesi Osman Gürdal
Süleyman Demirel Üniversitesi, Tıp Fakültesi Biyoistatistik
ve Tıbbi Bilişim Anabilim Dalı
32260 Çünür, Isparta



Üye : Prof. Dr. Soner ÇANKAYA
Ondokuz Mayıs Üniversitesi, Yaşar Doğu Spor Bilimleri
Fakültesi Spor Yöneticiliği Anabilim Dalı
55200 Atakum, Samsun



Üye : Dr. Öğr. Üyesi Adnan KARABRAHİMOĞLU
Süleyman Demirel Üniversitesi, Tıp Fakültesi Biyoistatistik
ve Tıbbi Bilişim Anabilim Dalı
32260 Çünür, Isparta



ONAY: Bu yüksek lisans tezi, Enstitü Yönetim Kurulu'nca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve kabul edilmiştir.

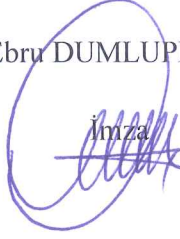
Doç Dr. Nilgün GÜRBÜZ
Enstitü Müdürü

BEYAN

“Klinik Vaka Tanısında ROC Analizi Tekniđi” adlı Yüksek Lisans tezi,
Süleyman Demirel Üniversitesi Sağlık Bilimleri Enstitüsü Lisansüstü Tez Önerisi ve
Tez Yazma Yönergesi 'ne uygun olarak hazırlanmıştır.

Tezi Hazırlayan

Ebru DUMLUPINAR

İmza


Danışman

Dr. Öğr. Üyesi Osman GÜRDAL

İmza



ÖNSÖZ

Tez konumun belirlenmesinde ve çalışmalarım süresince bilgi ve deneyimlerinden yararlandığım, yakın ilgi ve yardımlarıyla her zaman destek olan değerli hocam ve danışmanım sayın Dr. Öğr. Üyesi Osman GÜRDAL'a teşekkür ve saygılarımı sunarım.

Lisansüstü eğitim hayatımın başlamasında ve bu eğitim serüvenimin ilerleyen yıllarında desteğini, şefkatini ve anlayışını esirgemeyen değerli hocam Süleyman Demirel Üniversitesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı Başkanı sayın Prof. Dr. Hikmet ORHAN'a teşekkür ve saygılarımı sunarım.

Çalışmamda kullandığım verilerin sağlanmasında yardımını esirgemeyen Adıyaman Üniversitesi Eğitim ve Araştırma Hastanesi İç Hastalıkları Anabilim Dalı Başkanı Prof. Dr. Yusuf BİLEN'e ve Arş. Gör. Dr. Hatice ÖZDOĞAN'a teşekkür ve saygılarımı sunarım.

Bugüne kadar sabrını, şefkatini, sevgisini, özverisini hiç ama hiç esirgemeyen, benim ve kardeşimin mutluluğu, huzuru, istikrarı ve başarısı için hep yanımızda olan ve bugünlere gelmemi sağlayan annem Dudu ERARI'ya sonsuz saygı ve sevgilerimi sunarım.

Isparta, 2019



Canım anneme ithaf ediyorum...

Saygılarımla...

Isparta, 2019

İÇİNDEKİLER

KABUL ve ONAY SAYFASI	ii
BEYAN	iii
ÖNSÖZ	iv
İÇİNDEKİLER	vi
SİMGELER ve KISALTMALAR DİZİNİ	viii
TABLolar DİZİNİ	x
ŞEKİLLER DİZİNİ	xii
GRAFİKLER DİZİNİ	xiii
1. GİRİŞ	1
1.1. Araştırmanın Konusu	1
1.2. Araştırmanın Amacı	3
1.3. Araştırmanın Önemi	4
1.4. İlgili Araştırmalar	5
1.4.1. Receiver Operating Characteristic Tarihçesi	5
1.4.2. ROC Analizi ile İlgili Araştırmalar	5
2. GENEL BİLGİLER	11
2.1. Tanı Testleri ve Tıbbi Karar Oluşumu	11
2.2. Tanısal Test Değerlendirme Ölçütleri	14
2.2.1. 2x2 Atama Tablosu	17
2.2.2. Olabilirlik Oranı (Likelihood Ratio)	24
2.2.3. Bayes Teoremi	28
2.3. ROC Analizi	29
2.3.1. ROC Eğrisi	29
2.3.2. ROC Eğrisinin Değerlendirilmesinde Kullanılan Göstergeler.....	32
2.3.2.1. Eğri Altında Kalan Alan (AUC)	33
2.3.2.2. Tekil Noktalar ve Parçalı Alanlar	39
2.3.2.3. Youden İndeksi (YI)	39
2.3.2.4. En Yüksek Dikey Uzaklık (MVD).....	40
2.3.3. ROC Eğrisinin Kullanım Alanları.....	40
2.3.3.1. Optimum Eşik Değeri (Cutoff Value) Belirleme	40
2.3.3.2. Farklı Tanı Testlerinin Performanslarının Değerlendirilmesi.....	41

3. GEREÇ VE YÖNTEM.....	43
3.1. Araştırmanın Modeli	43
3.2. Çalışma Grubu	43
3.3. Verilerin Toplanması	44
3.4. Verilerin Analizi.....	44
3.4.1. Diyabet	44
3.4.2. Akut Pankreatit	46
4. BULGULAR	47
4.1. Diyabet	47
4.2. Akut Pankreatit.....	64
5. TARTIŞMA	74
6. SONUÇ VE ÖNERİLER.....	77
6.1. Sonuçlar.....	77
6.2. Öneriler	79
ÖZET.....	81
ABSTRACT	82
KAYNAKLAR	83
EKLER.....	86
Ek 1. MedCalc ve LABROC Ekran Görüntüleri	86
Ek 2. Etik Kurul İzni	90
Ek 3. Özgeçmiş	91

SİMGELER ve KISALTMALAR DİZİNİ

ADA	: Amerikan diyabet birliđi
APG	: Açlık plazma glukozu
AUC	: Eğri altında kalan alan
BT	: Bilgisayarlı tomografi
D⁻	: Tanı testi negatif olan
D⁺	: Tanı testi pozitif olan
DM	: Diabetes mellitus
FN	: Yanlış negatif
FNR	: Yanlış negatif oranı
FP	: Yanlış pozitif
FPR	: Yanlış pozitif oranı
GA	: Güven aralığı
HbA1c	: Hemoglobin A1c
LR⁻	: Negatif olabilirlik oranı
LR⁺	: Pozitif olabilirlik oranı
MVD	: En büyük dikey uzaklık
N	: Negatif
NDGG	: Amerikan ulusal diyabet veri grubu
NPV	: Negatif tahmin değeri
OGTT	: Şeker yükleme testi
p	: İstatistik p değeri
P	: Pozitif
pAUC	: Eğri altında kalan parçalı alan
PBM	: Uygun binormal model
PPV	: Pozitif tahmin değeri
ROC	: Receiver Operating Characteristic
S⁻	: Gerçekte sağlıklı olan
S. Hata	: Standart hata
S⁺	: Gerçekte hasta olan
TN	: Doğru negatif

TNR	: Doğru negatif oranı
TP	: Doğru pozitif
TPR	: Doğru pozitif oranı
var	: Varyans
WHO	: Dünya sağlık örgütü
x	: Örnek ortalaması
YI	: Youden indeksi
z	: z test istatistiği
ϕ	: Birikimli standart normal dağılım fonksiyonu
μ	: Popülasyon ortalaması



TABLolar DİZİNİ

Tablo 1. 1950-2007 Yılları Arasında Yapılan Yayın Sayısı	8
Tablo 2. 1966- 2017 Yılları Arasında ROC Analizi ile İlgili Yapılan Bazı Yayınlar.	9
Tablo 3. Farklı Alanlara Göre Değişken Adlandırmaları	16
Tablo 4. Kestirimci İçin İkili Doğru Atama Tablosu (Karar Matrisi).....	19
Tablo 5. Tanı Testi Değerlendirme Ölçütleri	28
Tablo 6. AUC Değerlerinin Yorumlanması	34
Tablo 7. Diyabet vakası cinsiyet dağılımları	44
Tablo 8. Akut pankreatit vakası cinsiyet dağılımları.....	44
Tablo 9. OGTT Sonucu Hasta ve Sağlıklı Bireylerin Dağılımları	47
Tablo 11. HbA1c Testinde 6,2 Eşik Değerine Göre İkili Atama Tablosu	49
Tablo 12. HbA1c Testi için ROC Eğrisi Altında Kalan Alan	50
Tablo 13. Parametrik Olmayan Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan ..	51
Tablo 14. Yarı Parametrik Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan.....	52
Tablo 15. Glukoz Testi için ROC Analizi Sonuçları.....	52
Tablo 16. Glukoz Testinde 116 Eşik Değerine Göre İkili Atama Tablosu	53
Tablo 17. Glukoz Testi için ROC Eğrisi Altında Kalan Alan.....	54
Tablo 18. Parametrik Olmayan Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan ..	55
Tablo 19. Yarı Parametrik Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan.....	55
Tablo 20. Üre Testi için ROC Analizi Sonuçları	56
Tablo 21. Üre Testinde 29 Eşik Değerine Göre İkili Atama Tablosu.....	57
Tablo 22. Üre Testi için ROC Eğrisi Altında Kalan Alan.....	58
Tablo 23. Parametrik Olmayan Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan ..	59
Tablo 24. Yarı Parametrik Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan.....	59
Tablo 25. Kreatinin Testi için ROC Analizi Sonuçları	60
Tablo 26. Kreatinin Testinde 0,8 Eşik Değerine Göre İkili Atama Tablosu.....	61
Tablo 27. Kreatinin Testi için ROC Eğrisi Altında Kalan Alan.....	62
Tablo 28. Parametrik Olmayan Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan ..	63
Tablo 29. Yarı Parametrik Yaklaşımına Göre ROC Eğrisi Altında Kalan Alan.....	63
Tablo 30. BT Sonucu Hasta ve Sağlıklı Bireylerin Dağılımları.....	65
Tablo 31. Lipaz Testi için ROC Analizi Sonuçları	65
Tablo 32. Lipaz Testinde 82 Eşik Değerine Göre İkili Atama Tablosu.....	66

Tablo 33. Lipaz Testi için ROC Eğrisi Altında Kalan Alan	67
Tablo 34. Parametrik Olmayan Yaklaşım Göre ROC Eğrisi Altında Kalan Alan ..	68
Tablo 35. Yarı Parametrik Yaklaşım Göre ROC Eğrisi Altında Kalan Alan	68
Tablo 36. Amilaz Testi için ROC Analizi Sonuçları.....	69
Tablo 37. Amilaz Testinde 90 Eşik Değerine Göre İkili Atama Tablosu	70
Tablo 38. Amilaz Testi için ROC Eğrisi Altında Kalan Alan.....	71
Tablo 39. Parametrik Olmayan Yaklaşım Göre ROC Eğrisi Altında Kalan Alan ..	72
Tablo 40. Yarı Parametrik Yaklaşım Göre ROC Eğrisi Altında Kalan Alan.....	72



ŞEKİLLER DİZİNİ

Şekil 1. Tanı Testlerinde Kesinlik ve Doğruluk Arasındaki İlişki.....	14
Şekil 2. Hasta ve Sağlıklı Olasılıkları	19
Şekil 3. Potansiyel Hasta ve Sağlıklı Olasılıkları	20
Şekil 4. Duyarlılık ile Yanlış Negatif Koşullu Olasılıkları.....	22
Şekil 5. Seçicilik ile Yanlış Pozitif Koşullu Olasılıkları	22
Şekil 6. Test Olasılıkları ile Hastalık Olasılığı Arasındaki İlişki.....	26
Şekil 7. Fagan Nomogramı	27
Şekil 8. ROC Eğrisi Örneği	31
Şekil 9. İki Test için ROC Eğrileri.....	31
Şekil 10. Test Sonuçlarının Kestirim Gücü ile ROC Eğrisi Arasındaki İlişki.....	32
Şekil 11. ROC Analizinde Değerlendirme Göstergeleri.....	33
Şekil 12. Parametrik Olan Yaklaşımına Göre Kestirilmiş ROC Eğrisi	36
Şekil 13. Parametrik Olmayan Yaklaşımına Göre Kestirilmiş ROC Eğrisi	37

GRAFİKLER DİZİNİ

Grafik 1. HbA1c Testi için ROC Eğrisi (%95 CI)	50
Grafik 2. Glukoz Testi için ROC Eğrisi (%95 CI).....	54
Grafik 3. Üre Testi için ROC Eğrisi (%95 CI).....	58
Grafik 4. Kreatinin Testi için ROC Eğrisi (%95 CI)	62
Grafik 5. Diyabet için Tanı Testlerinin Performans Karşılaştırmaları	64
Grafik 6. Lipaz Testi için ROC Eğrisi (%95 CI).....	67
Grafik 7. Amilaz Testi için ROC Eğrisi (%95 CI).....	71
Grafik 8. Akut Pankreatit için Tanı Testlerinin Performans Karşılaştırmaları	73

1. GİRİŞ

1.1. Araştırmanın Konusu

Geçmişten günümüze yaşam içerisinde birçok konuda ve alanda analiz edilmesi gereken problemler, ulaşılması gereken amaçlar söz konusudur. Yaşam içerisindeki problemlerin çözümlenmesi karar verme davranışı ile giderilebilecek bir süreçtir. Karar verme davranışı, var olan problem içerisindeki mevcut durumların belirli bir çözümleme mekanizması içerisinde değerlendirilmesi sonucu en uygun olan durumu seçebilmesidir. Değerlendirme sonucunda en uygun durum, temel alınan ya da öncesinde belli olan bir ölçüte göre belirlenir. Ölçüt, ya önceden kesin olarak belirlenmiş bir standarttır ya da gruba bağlı olarak çıkarılmış tipik bir puandır, normdur (1).

Sınıflandırma kavramını temel anlamda tanımlamak gerekirse, var olan dizindeki nesnelere ya da verilerin bir algoritma üzerinden tanımlı olan çeşitli kümelere ya da sınıflara dağıtmak olarak belirtilebilir. Problem veya sorun içerisindeki nesnelere ya da veriler bir atama prosedürü sonucunda hakkında gözlemlenen bilgiler temelinde bir sınıfa atanır. Bazı alınan her atama prosedürü mükemmel sonuçlar veren algoritmalara sahip olmayabilir. Bir nesnenin temel alınan atama prosedüründen dolayı yanlış bir sınıfa atanması oluşabilecek hatalardan birine örnektir. Bu tür hatalara sebep olan kusurlardan dolayı atama prosedürlerinin performans değerlendirilmeleri yapılması gerekmektedir. Prosedürlerin performans değerlendirilmeleri, belirlenen amaç doğrultusunda karar vermede yetkin olması için prosedürün geliştirilmesi ya da daha iyi bir atama prosedürü ile değiştirilmesi kararlarını almak için önemli bir işlemdir.

Yukarıda bahsedilen konuya uyan gerçek problemler ve amaçlara örnekler aşağıda verilmiştir (2).

1. Tıbbi teşhis yapmak

Amacımız, canlıdan elde edilen bilgiler neticesinde hastalık X veya hastalık Y'ye atamak

2. Konuşma sistemlerinin geliştirmek
Amacımız, konuşulan kelimeleri sınıflandırmak
3. Gen ifadesini içeren DNA çipi veri desenlerini araştırmak
Amacımız, DNA veri desenlerinin kansere uygun olup olmadığına karar vermek
4. Kredi kartı işlemlerini incelemek,
Amacımız, kredi kartının sahte olup olmadığına karar vermek
5. Gelen e-posta mesajlarını filtrelemek
Amacımız, e-postaların gerçek ya da spam olup olmadığına karar vermek
6. Üniversite öğrencilerini bir ders için değerlendirmek
Amacımız, öğrencilerin dersten geçip geçemeyeceğine karar vermek

Gündelik yaşamda istatistiksel kestirimlere pek çok alanda başvurulmaktadır. Sağlık, ekonomi, teknoloji, ulaşım ve eğitim gibi birçok yaşamsal faaliyetler üzerinde etkili olan kestirimlerin doğruluk oranına ilişkin hata payını bilmek önemlidir. Örneğin kredi başvurusu yapan kişi için kredi verme kararında ve uygulanacak faiz oranının belirlenmesinde borç ödeme olasılığının doğru olarak kestirilmesi istatistiksel kestirimciiler kullanılarak mümkün olabilmektedir. İstatistiksel kestirimlerden sıklıkla yararlanan diğer bir alan tıbbi karar verme sürecidir. Belirli şikâyetlerle hastalık belirtisi gösteren kişiler için konulan tıbbi tanı daha sonraki değerlendirmeler ve tedaviler üzerinde etkili olacağından, tanının doğruluk düzeyinin saptanması önemlidir. Tanı koymada maliyete ve hastanın taşıdığı riske göre değişen genellikle birden fazla test kullanılır ve elde edilen sonuçların etkinliğinin karşılaştırılmasıyla tanıya ilişkin değerlendirmeler yapılır. Örneklerde de belirtildiği gibi gelecek planları tüm sektörlerde, istatistiksel analizlerle elde edilen kestirimlere göre şekillenmekte ve kestirimlerin doğruluk düzeyini bilmek atılacak her adımda yol gösterici olmaktadır.

Tıbbi teşhis sürecinde doğru teşhisin yapılması, teşhis işleminin çok yönlü olmasından dolayı önemli bir problemdir. Var olan vakalar, iç ve dış kaynaklı nedenlerin bütününden dolayı kompleks bir yapıya dönüşmektedir. Tıbbi olgulardaki bilinmeyen nedenler teşhis sürecinde kesin yargılara ulaşmak yerine bazı olasılıkların belirlenmesine sebep olur (3).

Tıp alanında, klinisyenlerin sağlam karar vermeleri için tanı yöntemlerinin etkinlikleri tıbbi karar oluşumunda büyük öneme sahiptir ve tanı testlerinin etkinliklerinin değerlendirilmesi bir gereklilik haline gelmiştir. Doğru tanı koyma ve hasta birey ile sağlıklı bireyi ayırma yöntemleri sağlık alanında her zaman önem teşkil eden, ilgi duyulan ve de araştırılan bir konudur. Klinisyenler, bu araştırmalar neticesinde geliştirilen yeni yöntemlerin günden güne çoğalmasıyla var olan durumlarda hangi yöntemlere güvenilmeli ve bu yöntemlerin hangisi kullanılmalı sorularına cevap ararlar. Çünkü bir hastalığın tedavi sürecinde en önemli faktörün doğru tanı koymak olduğunu bilirler. Doğru tedavi tekniğine karar vermede önemli bir faktör olan tanıları yeterli güven düzeyinde olmalarının yanı sıra risksiz, öngörüsü yüksek ve de maliyetsiz olması büyük önem taşımaktadır.

Son yıllardaki yapılan çalışmalara ve gelişmelere bakıldığında tanı performansını değerlendirmede doğruluğun yanı sıra tanı testlerinin duyarlılığı ve özgüllüğü arasındaki ilişkiyi tanımlayacak analiz tekniklerinin uygulanması gerektiği belirtilmektedir. Yukarıda bahsedilen sebeplerden dolayı bilgisayar destekli karar verme mekanizmalarına dayalı uygulamalar temel alınarak, yapılan çalışmaların artış gösterdiği görülmektedir.

1.2. Araştırmanın Amacı

Tezin birincil amacı, tanı değerlendirme ölçütleri ile *Receiver Operating Characteristic* (ROC) analizini tanıtmak, kullanımı hakkında bilgi vermek ve anlatılan, verilen bilgilerin klinik vakalar üzerinden farklı istatistiksel programlar ile bir uygulamasını yapmaktır. Bu anlatımlarla birlikte, tıbbi araştırmacıların etkili bir şekilde kullanmasına yardımcı olmak için ROC analizinin temel kavramsal çerçevesini ve yorumunu sağlamaktır.

İkincil amaç ise, çalışmada yer verilen diyabet ve akut pankreatit vakalarında kullanılan tanı testlerinin performanslarını ve hastalık durumlarının en yüksek doğru sınıflama oranları ile ayırt edebileceği eşik değerlerini (kesim noktalarını) belirlemektir. Aynı zamanda, alt gruplar bazında tanı testlerinin performanslarının ve

eşik deęerlerinin deęişimini inceleyip, tanı testlerinin sonuçlarının daha etkin yorumlanmasını ve de kullanılması sağlamaktır.

1.3. Araştırmanın Önemi

Saęlık alanında klinisyenler ve araştırmacılar eksik bilgi ve belirsiz sonuçlara rağmen saęlam kararlar vermek için tanı testlerine başvururlar. Bu testler, erken ve doğru teşhisin konması açısından önemli bir yere sahiptir. Belli başlı bazı vakalarda saęlıklı ve saęlıksız kesin ve doğru ayıran testler bulunmaktadır. Bu testlerin, çoęu zaman hasta üzerinde riskler oluşturan, zaman alıcı, hastane giderleri açısından maliyeti yüksek olan uygulaması zor testler olmasından dolayı kullanılması sürekli mümkün olmayabilir. Yapılan araştırmalar ile hastalık durumunu kesin ve doğru şekilde ayırabilen testlere, daha az risksiz, daha az zaman alıcı, daha maliyetsiz ve de kestirim gücü mükemmel yakın alternatif olabilecek tanı testleri geliştirilmeye çalışılmaktadır. Geliştirilen alternatif testler belirtilen olumlu yönleri nedeniyle tıbbi karar oluşumunda tercih edilebilir. Bu testler tıbbi karar oluşumunda kullanılmadan önce performans deęerlendirmelerinin yapılması ve hastalık durumunu en yüksek doğru sınıflama yapacak optimum eşik deęerinin bulunması gerekir. Bu deęerlendirmelerin ve işlemlerin yapılması açısından ROC analizi önemli bir istatistiksel tekniktir.

1.4. İlgili Arařtırmalar

1.4.1. Receiver Operating Characteristic Tarihçesi

Alıcı iřletme karakteristik analizi, bařlangıçta II. Dünya savařı sırasında radar tespitindeki sinyallerden gelen gürültüyü ayırt etmedeki sınıflandırma doęruluęunu analiz etmek için geliřtirilmiřtir (4). İstatistiksel karar teorisine dayanan ROC analizi 1950'lerin bařında teknik bilimlerde sinyal algılama analizi için kullanılmaya bařlanmış, ilerleyen dönemlerde temel alanlarda kullanılan istatistiksel tekniklerden birisi olmuřtur. Daha sonraki senelerde ilk kez psikofizik alanındaki arařtırmalarda uygulanmıřtır. 1960 yılının bařlarında tıp alanında da kullanılabileceęi fikri ortaya atılmıř, 1960 yılının sonlarına doęru görüntüleme cihazlarının deęerlendirilmesinde kullanılmıřtır. ROC analizinin tıbbi karar vermede bu řekilde yer bulmasını saęlayan ve bu rolü tařıyan ilk kiři Dr. Lee Lusted olmuřtur. Daha sonra, 1970'lerde ve 1980'lerde teknięin tıbbi test deęerlendirme ve karar verme ile ilgili olarak önemli olduęu ve o zamandan bu yana geçen yıllarda teknięin radyoloji, kardiyoloji gibi alanlarda çok fazla geliřme ve kullanım görmüřtür. ROC analizi ilk zamanlarda daha geniř bir uygulama yelpazesine girmeden benzer alanlarda kullanılmıřtır. Yakın geçmiřten bu yana sosyoloji, deneysel psikoloji, atmosferik ve yer bilimleri, finans, makine öğrenimi ve veri madencilięi gibi çeřitli konularda yer bulmuřtur. Son zamanlarda ise ROC metodolojisi yoęun olarak tarama ve tanı testlerine baęımlı olan klinik alanlara adapte edilmiřtir. Özellikle laboratuvar testlerinin etkinlikleri, epidemiyoloji ve biyoinformatik alanları ile iliřkilendirilmektedir.

1.4.2. ROC Analizi ile İlgili Arařtırmalar

ROC analizi ile ilgili yapılan literatür taramasında ROC analizinin ilk kullanılmaya bařlandığı 1950 ve 1960 yıllardan günümüze kadar farklı alanlarda yer bulduęu görölmüřtür. Bu literatür taramasının bir özet deęerlendirmesi olarak geçmiřte baz aldıęı konular sebebiyle önemli yer edinmiř ve řimdiki çalıřmalarda

başvurulan kaynaklara ve yakın geçmişte tıp alanında ROC analizinin kullanıldığı bazı çalışmalara yer verilecektir.

Green ve Swets'in 1966 yılında yayınladıkları “Signal Detection Theory and Psychophysics” adlı yazında istatistiksel karar verme teorisini ve bunun çeşitli ortak psikofiziksel faaliyetlere uygulanmasını ana hatlarıyla belirtir. Sinyal algılama teorisini duyuşal psikolojideki problemlere uygular (5).

Lusted'in 1971 yılında yayınladığı “Signal Detectability and Medical Decision Making” adlı yayınında tıbbi karar vermede ROC analizinin kullanılabilirliğini öne sürmüş ve radyolojide kullanılan farklı tekniklerin tanı tutarlılıkları üzerindeki performans değerlendirmelerini işlemiştir. ROC analizine tıbbi karar vermede ilk rol bu çalışmada verilmiştir (4).

Goodenough, Rossmann ve Lusted'in 1972 yılında yayınlanmış olan “Radiographic Applications of Signal Detection Theory” isimli çalışmalarında ROC eğrilerinin, radyografik bir arka planda oluşan sinyallerin tespitinde gözlemci performansını tanımladığını ve daha fazla sinyal algılanabilirliğini sağladığını ifade etmiştir (6).

Egan'in 1975 yılında yayınladığı “Signal Detection Theory and ROC Analysis” adlı yayınında alıcı çalışma karakteristik grafiğinin performanslara göre sınıflandırıcıları görselleştirmek, organize etmek ve seçmek için kullanılan bir teknik olduğunu ve ROC grafiklerinin sinyal algılama teorisinde, sınıflandırıcıların isabet oranları ve yanlış alarm oranları arasındaki deęiş tokuşu göstermek için kullanıldığını ifade etmiştir (7).

Weinstein Obuchowski ve Lieber'in 2005 yılında yaptıkları “Clinical Evaluation of Diagnostic Tests” adlı çalışmada, tanı test doğruluğunun temel tanımlarını irdelenmişler, ardından, “Bu hasta bu tanı testine tabi mi olmalı?” sorusunu ele almak için tanı testlerinin nasıl kullanılabileceğini değerlendirmiş ve test sonucundan sonra, “Bu hastanın hastalığı olma olasılığı nedir?” sorularına yanıt bulmak için tasarlanan tanısal test doğruluğunu tahmin eden veya karşılaştıran araştırmalar içerisindeki kavramlar üzerine tartışmışlardır (8).

Zweig ve Campbell'in 1993 yılında yayınlamış oldukları “Receiver Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine” adlı çalışmalarında bir laboratuvar testinin klinik performansının, teşhis doğruluğunun klinik olarak ilgili alt gruplara doğru şekilde sınıflandırma kabiliyeti ile tanımlanabildiğini ve teşhis doğruluğunun, sınıflandırma cihazı tarafından sağlanan bilgilerin kalitesi ile ifade edildiğini belirtmiştir. ROC grafiklerinin, tüm çalışma koşullarının spektrumunda alternatif sağlık durumları arasında ayırım yapabilme yeteneğinin sınırlarını gösteren düzgün bir doğruluk indeksini sağladığını öne sürmüşlerdir (9).

Metz'in 1978 yılında yaptığı “Basic Principles of ROC Analysis” isimli çalışmasında alıcı çalışma karakteristiği eğrisinde karar eşik etkisinin basit fakat tam bir ampirik açıklaması olduğunu, çeşitli doğru ve yanlış kararların göreceli frekanslarının tüm olası kombinasyonlarını gösteren kesim noktalarına sahip olduğunu belirtmiştir. Aynı zamanda ROC analizinin tanısal karar vermenin maliyet-fayda analiziyle doğrudan ve doğal bir şekilde ilişkili olduğunu ifade etmiştir (10).

Hanley ve McNeil'in 1982 yılında yayınlanan “The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve” isimli çalışmalarında tanı testlerinin performans değerlendirme için ROC analizinde yer alan göstergelerden birisi olan eğri altında kalan (AUC) değerinden yararlanılabileceği öne sürülmüştür (11).

Hanley ve McNeil'in 1983 yılında yayınladıkları “A Method of Comparing the Areas under Receiver Operating Characteristic (ROC) Curves Derived from the Same Cases” isimli diğer çalışmada ise tanı testlerinin performanslarını AUC değerlerinin kıyaslaması ile değerlendirilirken korelasyonunda ele alınması gerektiğini ifade etmişlerdir (12).

Kim, Lim ve Oh'un 2004 yılında çalıştıkları “Preoperative MRI of Rectal Cancer with and without Rectal Water Filling: an Intraindividual Comparison” isimli çalışmada ılık su kullanılarak yapılan rektal distansiyonun, rektal kanserin preoperatif evrelemesi için MRG'nin doğruluğunu artırıp arttırmadığını araştırmışlardır. Bitişik yapıların invazyonu ve bölgesel lenf nodu, tutulumu

açısından altın standart kabul edilmiştir. Çalışmanın sonucunda dış duvar penetrasyonunun belirlenmesinde tipik gözden geçiricilerin doğruluğunun rektum distansiyonu ile arttığını, ancak bölgesel lenf nodu tutulumunun belirlenmesinde gözden geçirenlerin doğruluğunun etkilenmediğini ifade etmişlerdir (13).

Osada, Kaku, Litsuka, Seki ve Sekiya'nın 2003 yılında yayınladıkları “Quantitative and Qualitative Evaluations of Fetal Lung with MR Imaging” adlı çalışmada MRG'nin fetal pulmoner hipoplaziyi öngörme yeteneğini değerlendirmek için ROC analizini kullanmışlardır. Akciğer hacmine dayanan ROC eğrisinin, akciğer hacminin doğum sonrası zayıf solunum sonucu olanlara karşı iyi olan fetüsler arasında ayırım yapma yeteneğine sahip olduğunu belirtmişlerdir. Bununla birlikte, akciğer hacminden ve sinyal yoğunluğundan gelen birleşik bilgilere dayanan ROC eğrisinin daha yüksek hassasiyete sahip olduğunu ileri sürmüşlerdir (14).

Zheng, Leader ve Ambrams'ın 2004 yılında yaptıkları “Computer-aided Detection Schemes: The Effect of Limiting the Number of Cued Regions in Each Case” adlı araştırmada mamografik bilgisayar destekli algılama (CAD) planının doğruluğunun pozitif olarak tanımlanabilecek maksimum bölge sayısını kısıtlayarak nasıl etkilendiğini değerlendirmişlerdir. Serbest-yanıtlı alıcı işletim karakteristik eğrileri (FROC) adı verilen ROC eğrisinden farklı özel bir ROC tekniği kullanmışlardır (15).

ROC analizi ile ilgili çalışmalar, 1960 ve 1980 yılları arasında minimum seviyede, ROC eğrisinin farklı disiplinlerde kullanılmaya başlamasından itibaren yapılan araştırmaların sayısında bir artış gözlenmiştir. ROC analizi ile ilgili yayınların sayıları bazı yıllar bazında Tablo 1’de verilmiştir.

Tablo 1. 1950-2007 Yılları Arasında Yapılan Yayın Sayısı

Tarihler	Yayın Sayısı
1964 ve öncesi	2
1964 – 1979	42
1980 – 1987	70
1988 – 1995	1046
1996 - 2003	4088
2004 - 2007	4463

Tablo 2. 1966- 2017 Yılları Arasında ROC Analizi ile İlgili Yapılan Bazı Yayınlar

Yazar	Yıl	Yayın
DM Green ve JA Swets	1966	Signal Detection Theory and Psychophysics
DD Dorfman ve Edward JR Alf	1968	Maximum Likelihood Estimation of Parameters of Signal Detection Theory
LB Lusted	1971	Signal Detectability and Medical Decision Making
JP Egan	1975	Signal Detection Theory and ROC Analysis
D Bamber	1975	The Area Above the Ordinal Dominance Graph and the Area Below the Receiver Operating Characteristic Graph
CE Metz	1978	Basic Principles of ROC Analysis
JA Hanley ve BJ McNeil	1982	The Meaning and Use of the Area Under a Receiver Operating Characteristic(ROC) Curve
JA Swets ve RM Pickett	1982	Evaluation of Diagnostic Systems: Methods from Signal Detection Theory
JA Hanley ve BJ McNeil	1983	A Method of Comparing the Areas under Receiver Operating Characteristic (ROC) Curves Derived from the Same Cases
CE Metz	1986	ROC Methodology in Radiology Imaging
AN Tosteson ve CB Begg	1988	A General Regression Methodology for ROC Curve Estimation
ER DeLong, DM DeLong ve DL Clarke-Pearson	1988	Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach
RF Raubertas, RE Rodewald, SG Humiston ve PG Szilagyi	1994	ROC Curves for Classification Trees
F Hsieh ve BW Turnbull	1996	Nonparametric and Semiparametric Estimation of the Receiver Operating Characteristic Curve
M Coffin ve S Sukhatme,	1997	Receiver Operating Characteristic Studies and Measurement Errors
NA Obuchowski. ve DK McClish	1997	Sample Size Determination for Diagnostic Accuracy Studies Involving Binormal ROC Curve Parameters
G Li, RC Tiwari ve MT Wells	1999	Semiparametric Inference for a Quantile Comparison Function with Applications to Receiver Operating Characteristic Curves

Tablo 2. (Devamı) 1966- 2008 Yılları Arasında ROC Analizi ile İlgili Yapılan Bazı Yayınlar

Yazar	Yıl	Yayın
KH Zou ve WJ Hall	2000	Two Transformation Models for Estimating an ROC Curve Derived from Continuous Data
D Faraggi ve B Reiser	2002	Estimation of the Area under The ROC Curve
KO Hajian-Tilaki ve JA Hanley	2002	Comparison of Three Methods for Estimating the Standard Error of the Area Under the Curve in ROC Analysis of Quantitative Data
J Qin ve B Zhang,	2003	Using Logistic Regression Procedures for Estimating Receiver Operating Characteristic Curves
Z Zhang	2004	Least Squares Analysis of the Receiver Operating Characteristic Curve
T Fawcett	2004	ROC Graphs: Notes and Practical Considerations for Researchers
CT Nakas ve CT Yiannoutsos	2004	Ordered Multiple-Class ROC Analysis with Continuous Measurements
M. Gönen	2007	Analyzing Receiver Operating Characteristic Curves Using SAS
CE Metz	2008	ROC Analysis in Medical Imaging: A Tutorial Review of the Literature
R Kumar, A Indrayan	2011	Receiver operating characteristic (ROC) curve for medical researchers.
K Hajian-Tilaki	2013	Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation
AI Bandos, B Guo, D Gur	2017	Jackknife variance of the partial area under the empirical receiver operating characteristic curve.

2. GENEL BİLGİLER

2.1. Tanı Testleri ve Tıbbi Karar Oluşumu

Klinisyenin başlıca görevi, eksik klinik bilgi ve klinik sonuç belirsizliğine rağmen hasta bakımı konusunda sağlam kararlar vermektir. Hastanın hikayesi ve fizik muayene ile elde edilen veriler genellikle klinik karar ve tedavinin yönlendirilişi için yeterli olsa da daha fazla bilgi gerekebilir. Böyle durumlarda klinisyenler genellikle tanı testlerine başvururlar.

Hasta yönetimi ile ilgili soruların cevaplanmasına yardımcı olmak için laboratuvar testleri istenir. Bir test sonucunun ne kadar yardım sağladığını ve her durumda çok karmaşık bir konudur. Yönetim kararları ve stratejileri, hekimin hastalık olasılığını, mevcut verilerin kalitesini, çeşitli tedavi/yönetim alternatiflerinin etkinliğini, sonuçların olasılığını, sonuçların maliyetini dikkate almasını gerektiren karmaşık faaliyetlerdir. Çoğu zaman klinik veriler genellikle karmaşık bir karar sürecine entegre edilir. Tek bir laboratuvar test sonucu, bir teşhis veya hasta yönetimi kararının tek temeli değildir (16).

Eğer doğru kullanılırsa tanısal testler klinisyene büyük yarar sağlar. Testler tarama, yani hastalık risk faktörlerini tanımlamak ve semptomatik kişilerde gizli hastalığı saptamak amacıyla yapılabilir. Risk faktörlerinin saptanması hastalığı önlemeye yönelik erken girişime olanak sağlar, gizli bir hastalığın erken saptanmasına erken tedavi ve dolayısıyla hastalık morbidite ve mortalitesini azaltabilir. Testler tanı, yani semptomatik kişilerde hastalığı saptama ya da dışlama amacıyla kullanılabilir. Bazı testler semptom ve belirtilerin başlamasından sonra erken tanıya, bazıları ayırıcı tanıya, bazılarıysa hastalık evresi ya da aktivitesinin belirlenmesine yardımcı olur (16).

Testler hasta yönetiminde aşağıdaki durumlar noktasında yararlı olabilir:

1. Hastalık şiddetini belirlemek
2. Prognozu tahmin etmek
3. Hastalığın gidişatını izlemek

4. Hastalığın nüksünü tahmin etmek
5. İlaç seçimi ve tedavinin düzenlenmesi

Bununla birlikte her klinik araç; bir klinik kimya testi, bir elektroensefalogram, bir elektrokardiyogram, bir nükleit taraması, bir röntgenogram, bir biyopsi, bir delikten bir görünüm, bir solunum fonksiyon testi veya bir sonogram hastalığın olup olmadığına dair bazı belirgin ayrımları yapmak içindir.

Medikal arařtırmalarda karar alma işlemleri daha karmaşık problemlerin çözümünde arařtırmacıya kolaylıklar sunan bir metotlar bütünlüğü olarak bilinmektedir. Bu bütün içerisinde hastalığın teşhisi, uygulanacak olan tedavi türünün belirlenmesi, tedavide kullanılacak olan ilaçların seçimi ve tedavi yolundaki her türlü maddi giderler ayrı ayrı incelenmesi gereken olgulardır.

Medikal arařtırmacılar doktorlar ve hastane yönetimiyle ilgili yönetimde yer alan herkes çalışmalarının çeşitli evrelerinde kararlar alırlar. Medikal arařtırmacılar, arařtırma konusu doğrultusunda yapacakları çalışmalarda karşılaştıkları çeşitli sonuçlardan tek sonuca varmak için doktorlar hastalıkların teşhisinde ve bu teşhise göre uygulanacak tedavi yöntemlerinin seçiminde hastane yönetimi de yapılan her arařtırmanın ve uygulanan tanıların maliyetlerini de göz önüne alarak hastane olanaklarının gelişiminde çeşitli kararlar alırlar. Bu tür konularda karar yetkisine sahip kişilerin alacakları kararlar kimi zaman çıkacak bir ilacın güvenilirliğini, kimi zaman hastaların hayatlarını, kimi zamanda hastane olanaklarının en uygun seviyede kalıp kalmayacağını etkileyen önemli kararlar olmaktadır (17).

Tanısal performansın herhangi bir değerlendirmesi, tanısal kararların "gerçek" ile karşılaştırılmasını gerektirir. Muhtemelen en basit teşhis karar kalitesi ölçütü, hekimin doğru olduğu, genellikle "doğruluk" olarak adlandırılan vakaların kesridir. Her ne kadar hepimiz yüksek doğruluğun iyi olduğunu kabul etmeye istekli olsak da sayı çok yanıltıcı olabilir. Örneğin, nispeten nadir görülen bir hastalığın taranmasında, tüm kanıtlar görmezden gelinerek ve tüm vakaları olumsuz olarak adlandırmak çok kesin olabilir. Hastaların sadece %5'inde söz konusu hastalık varsa, her zaman kör bir şekilde hastalığın bulunmadığını belirten bir doktor % 95'inin doğru olacağını söyler. Doğruluk, bir teşhis performans endeksi olarak sınırlı bir

fayda sağlar çünkü hastalık prevalansı sonuçtaki sayıyı güçlü bir şekilde etkiler ve hastalık prevalansı için hiçbir matematiksel düzeltme bu endeksi anlamlı şekilde kullanamaz. Bu doğru olsa da doğruluğun anlamlı olması gerektiğini varsaymak cazip gelebilir. Gerçekte bunun istatistiksel olarak analiz edilmesi ve kanıtlanması gerekmektedir.

Diğer taraftan yanlış pozitif bir test sonucu yanlış tanıya ya da gereksiz daha ileri testlere kapı açabilir. Sağlıklı bir kişiyi yanlış pozitif bir test sonucunda hasta göstermek psikolojik etkilenme yanı sıra gereksiz ya da uygunsuz tedavi riskine yol açabilir. Aslında bu istatistiksel tanıda I. hata oranı kabul edilmiş anlamındadır.

Tanısal testlerin maliyeti de göz önüne alınmalıdır. Toplam maliyet yüksek olabilir veya maliyet bakımından yetersiz kalabilir. Nispeten, pahalı olmayan testlerin bile, eğer sağlığa katkısı küçükse, maliyet açısından yararı olmayabilir.

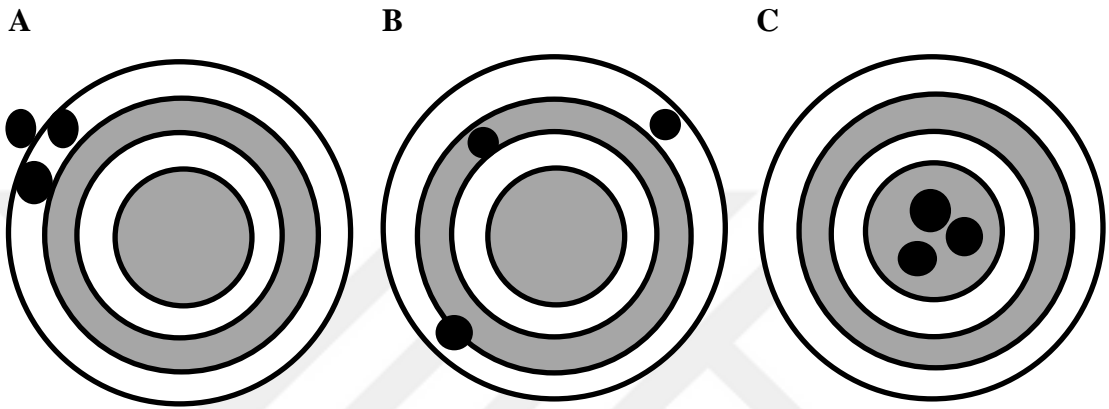
Klinisyenler her gün çok sayıda laboratuvar testi isterler ve yorumlarlar, istenilen bu testlerin karmaşıklığı giderek artmaktadır. Büyük ve gittikçe genişleyen test menüsü doğru laboratuvar testi seçiminde ve test sonuçlarının doğru yorumlanmasında klinisyenleri zorlamaktadır. Test seçimi ve yorumundaki hatalar sık yapılır ama saptamak zordur. Spesifik bozukluklarda test seçimine kılavuzluk yapan kanıta dayalı test algoritmaları ve uzmanlarca yapılan test yorumu bu tür hataları azaltabilir, zamanında ve doğru tanıyı geliştirir.

Aşağıda yararlı tanı testlerinin genel özellikleri listelenmiştir. Aşağıda ayrıntısıyla verilen ilkeler sadece laboratuvar testleri ve radyolojik testler değil aynı zamanda hastanın hikayesi ve fizik muayene bilgileri içinde geçerlidir. Bu özelliklerin anlaşılması tanı testlerini isterken ve yorumlarken klinisyene çok yardımcı olur (16).

1. Test metodolojisi ayrıntılı tanımlanmalıdır, böylece test doğru ve güvenilir şekilde tekrarlanabilir.
2. Test doğruluğu ve kesinliği belirlenmelidir.
3. Referans aralığı uygun şekilde belirlenmelidir.

4. Duyarlılık ve özgüllük bir altın standart test ile karşılaştırılarak güvenilir şekilde tespit edilmelidir. Bu karşılaştırmada farklı ama sıklıkla birbirine karışan bozuklukları olan, spektrumu orta derecede veya şiddetlidir.

Aşağıda tanısal testlerde doğruluk ve kesinlik arasındaki ilişki, hedefin merkezi test edilen maddenin gerçek değerini gösterdiği varsayımıyla, şematik şekilde gösterilmiştir.



Şekil 1. Tanı Testlerinde Kesinlik ve Doğruluk Arasındaki İlişki

A şekli, kesin ama doğru olmayan bir tanı testini ifade etmektedir. Tekrarlanan ölçümler çok yakın sonuçlar vermiş olup, sonuçlar hedef değere uzaktır.

B şekli, kesinliği ve doğruluğu kötü bir tanı testini ifade etmektedir. Tekrarlanan ölçümler çok farklı sonuçlar vermiş olup, sonuçlar hedef değere uzaktır.

C şekli, hem kesinliği hem doğruluğu iyi olan ideal bir tanı testidir.

2.2. Tanısal Test Değerlendirme Ölçütleri

Klinik karar almalar, sağlık alanında oluşan problemlerin çözümlenmesinde sıklıkla kullanılan yöntemlerdir. Var olan problemlere bu yöntemler yardımıyla olasılık kurallarına göre kıyaslanarak çözümler bulunmaktadır (18).

Klinik karar alma ile ilgili problemlerin çözümlerinde, tanı testlerinin doğruluğunun değerlendirilmesi, test sonuçlarının baz alınan hastaya göre

yorumlanması, kompleks hasta yapısından kaynaklı problemlerin çözümlenmesi ve problemin çözümünde kullanılacak en iyi ölçütün ya da metodun seçilmesi ele alınan önemli noktalar. Kanıta dayalı tıbbın uygulamasında ve kullanılmasında önemli yer teşkil etmektedir (19).

Klinik karar almada kullanılan bir tanı testinin doğruluğu ya da kesinliği; doğru sonuç verip vermemesi yani gerçekte sağlıklı ve hasta olan bireyleri test sonucuna göre aynı durumlara atama etkinliğidir.

Tanı testlerinin doğruluklarını değerlendirmek için birçok ölçüt geliştirilmiştir. Tanı testlerinin sonuçlarının türlerine göre farklı değerlendirme ölçütleri kullanılır. Bir tanı testi sonucuna göre sürekli, dikotom (ikili) ve ordinal (sıralı) olarak adlandırılır. Örneğin; kan şekeri sonucunu veren bir laboratuvar testi sürekli, pankreas iltihabı olup olmadığını veren radyolojik görüntüleme dikotom ve kanseri evrelerine (evre I, evre II, evre III) göre veren testte ordinal olarak adlandırılabilir. En sık karşılaşılan tanı testleri sürekli ve dikotom sonuç veren tanı testleridir. Tanı testi olarak kullanılan biyokimya laboratuvar ve radyolojik görüntüleme sonuçları da çoğunlukla dikotom ve sürekli sonuç veren testlerdir. Bu testlerin değerlendirilmesi için uygulanan istatistiksel analizde sıklıkla kullanılan ölçütler 2x2 tablolar, olabilirlik oranları ve de Bayes Teoremidir. Tanı testinin sonucunun sürekli, gerçek durumun ise dikotom olması durumunda ise tanı testinin performans değerlendirilmesinde ve tanı testinin optimal eşik (cutoff) değerinin belirlenmesinde ise ROC analizi yöntemi kullanılır (20-22).

Ele aldığımız konun farklı alanlarda uygulanmasına karşın ortak bir teması vardır. Öngörülen varlığın değeri, sonucu bilinmeden önce bir tahmin yapmak ya da kestirim yapmak konunun temelini oluşturur. İrdelenecek ve üzerinde durulacak kısım ise yapılan kestirimlerin ya da tahminlerin doğruluğunu değerlendirmektir. Bir problem üzerinde verilecek kararda hesaplanan olasılıkların doğru olarak tahmin edilmesi istatistiksel kestirimciler kullanılarak mümkün olabilmektedir. Sağlık alanında ise uygulanan tanı testlerinin doğruluklarının analiz edilmesi ve belirlenmesi hem hasta hem hekim hem de sağlık birimi için önem teşkil etmektedir. Bu tahminlerin doğruluk düzeylerinin değerlendirilmesinde kullanılan çeşitli

yöntemler ve istatistiksel analizler vardır. Yukarıda da bahsedildiği gibi bu yöntemlerden birisi de ROC analizidir.

Yukarıda bahsettiğimiz yapılan kestirimler veya tahminlerde, tahminlerin ve sonuçların isimleri araştırmanın alanına göre farklılık göstermektedir. Kestirimlerin doğruluklarının değerlendirilmesinde kullanılan ikili atama tablosunun kullanım alanına bağlı olarak en sık kullanılan farklı adlandırma ifadelerine Tablo 3’de yer verilmiştir.

Tablo 3. Farklı Alanlara Göre Değişken Adlandırmaları

Test Durumu Değişkeni	Gerçek Durum Değişkeni	Değişken Değerleri
Kestirimci	Sonuç	Gözlem - Kontrol
Skor/Puan	Altın Standart	Pozitif - Negatif
Tahmin	Gösterge	Var - Yok
Tanı/Teşhis	Gözlemlenen	İyi - Kötü
Durum	Sonuç	Normal - Anormal
Karar	Referans	Sağlam - Hasta

Yukarıda atama tablosunda farklı nitelendirmelere yer verilen gerçek durum ile test durumunu belirlemede kullanılan testlerin özellikleri ayrı ayrı aşağıdaki şekilde verilmiştir.

Altın standart (Referans test) olarak kullanılan testler;

1. Gerçek durumu ya da gerçek duruma en yakın sonucu yansıtır.
2. Hasta ve sağlıklı bireyleri birbirinden kesin ayırır veya ayırma gücü yüksektir.
3. Genelde riskli, maliyetli, zor ve zaman alıcıdır.

Tanı testi olarak kullanılan testler;

1. Gerçek durumu tam anlamıyla yansıtmazlar.
2. Hasta ve sağlıklı bireyleri birbirinden kesin olarak ayıramazlar.
3. Genelde risksiz, maliyetsiz, kolay ve de zaman almazlar.

Tanı testlerinin kestirimlerinin doğrulukları ya da kesinlikleri altın standart olarak alınan referans testi sonuçları ile belirlenir (22).

Tanı testlerinin doğruluklarının ya da kesinliklerinin değerlendirilmesinde sıklıkla kullanılan ve olasılık kuralları üzerine kurulan ölçütler olan 2x2 atama tabloları, olabilirlik oranları ve Bayes Teoremi sırasıyla incelenecektir. Kullanılabilir olduğu düşünülen tanı testinin klinik vakalardaki rolü belirlenmeli ve de kestirim performansı değerlendirilmelidir. Bu değerlendirmede testin sadece ayırt edicilik gücüne değil bununla birlikte hastalığın prevalansı, durumu ve yanlış tanılarının sonuçlarına yer verilmelidir.

2.2.1. 2x2 Atama Tablosu

Bu bölümde altın standart olarak temel alınan referans testine göre ROC analizinde kullanılan doğru pozitif kesir (duyarlılık), doğru negatif kesir (seçicilik), yanlış pozitif kesir, yanlış negatif kesir, pozitif öngörü değeri ve negatif öngörü değerinin teorik yapısı ve hesaplama şekillerine yer verilecektir.

Bir tanı testinin doğruluğu çeşitli yollarla tanımlanabilir. Duyarlılık (sensitivity) ve seçicilik (specificity) testin doğruluğunun değerlendirilmesinde kullanılan başlıca ölçütlerdir. Bu ölçütler en anlaşılır şekilde 2x2 atama tabloları üzerinden anlatılabilir.

Başlangıç noktamız olarak, etiketlenmiş bireyleri tahsis etmede var olan durum değişkeninin “pozitif” (P) ve “negatif” (N) olmak üzere iki popülasyonu kapsadığı varsayalım. Belirlenen bir sınıflama kuralıyla birlikte bireyler rassal olarak pozitif ve negatif popülasyonlar olarak iki sınıfa ayrılır. Bu sınıflandırma kuralının her bireyde ölçülen değişken vektörünün (X), sürekli fonksiyonunun $h(X)$ olduğu varsayılır. Böylece fonksiyonun büyük değeri pozitif popülasyonun göstergesidir ve küçük değerlerin negatif popülasyonun göstergesi olduğunu gösterir. Bu nedenle, eğer X belirli bir birey için X 'in gözlenen değeri ise ve $h(X)$ bu birey için fonksiyon skoru ise o zaman birey $h(X)$ 'e göre bir eşiği aştığı veya aşmadığı şekilde pozitif popülasyona veya negatif popülasyona tahsis edilir. Sınıflandırma fonksiyonu $h(X)$,

analiz belirleyici bileşeni olduğunda ilgili olasılık modelleri, başlangıçtaki X ölçümlerinin rastgele vektöründen elde edilen sınıflandırma puanları için olandır (2).

Bu bölümde sadece dikotom sonuç veren tanı testlerini ele alıyoruz. Oluşturduğumuz karar matrisinin sütununda altın standarda göre belirlenmiş gerçek durum, satırda ise tanı testinin sonucu yer alır. Bazı testler doğası gereği iki yönlüdür. Bir cevaplayıcı örneğin bir soruya “evet” veya “hayır” cevabını verir. Bazı testler ise sürekli sonuçlar verir ancak pratikte bir eşik değeri kullanılarak dikotomize edilir (23).

Gerçek hastalık durumunu göstermek için ikili değişken olan S 'yi kullanalım;

$$S = \left\{ \begin{array}{l} \boxed{1, \quad \text{Hastalık var } (S^+)} \\ \boxed{0, \quad \text{Hastalık yok } (S^-)} \end{array} \right\} \quad (3.1)$$

D değişkeni, tanı (teşhis) testinin sonucudur. Genel olarak daha büyük değerinin hastalığın göstergesi olduğu kabul edilir.

$$D = \left\{ \begin{array}{l} \boxed{1, \quad \text{Hastalık için pozitif } (D^+)} \\ \boxed{0, \quad \text{Hastalık için negatif } (D^-)} \end{array} \right\} \quad (3.2)$$

Testin sonucu aşağıda gösterildiği gibi gerçek doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif olarak sınıflandırılabilir. Nitelendirmelerden de anlaşılacağı gibi, hastalıklı bir kişi pozitif bir test sonucu ile doğru bir şekilde sınıflandırıldığında doğru pozitif (TP) sonuç, hastalıklı bir kişi negatif test sonucu ile sınıflandırıldığında ise yanlış negatif (FN) sonuç ortaya çıkar. Benzer şekilde doğru negatif (TN) ve yanlış pozitif (FP), deneğin sırasıyla negatif veya pozitif bir sonucu olduğunda ortaya çıkar (22, 23).

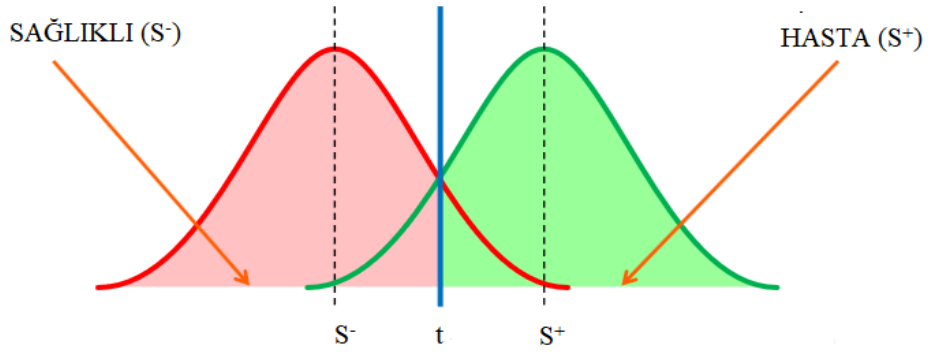
Yukarıda verdiğimiz tüm test sonuçlarını hastalık durumlarına göre Tablo 4' deki kontenjans tablosu ile sınıflandırılabilir ve özetleyebiliriz. Bu tablo var olan kestirimceiler için ikili doğru atama tablosudur. Doğru atama tabloları gerçek durum

ve kestirilen sınıf değerlerine dayalı olarak D^+ ve D^- sınıflarındaki her bir birim değerinin eşik değeri ile karşılaştırılmasıyla elde edilir.

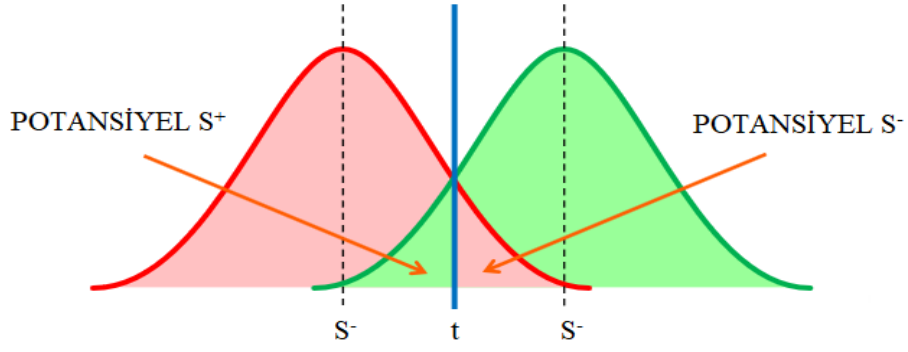
Tablo 4. Kestirimci İçin İkili Doğru Atama Tablosu (Karar Matrisi)

Tanı Testi Sonucu (Tahmin)	Gerçek Durum		
	Hastalık (S^+)	Sağlıklı (S^-)	Toplam
	Pozitif (D^+)	<i>TP</i> (Doğru pozitif) Duyarlılık a	<i>FP</i> (Yanlış pozitif) c
Negatif (D^-)	<i>FN</i> (Yanlış negatif) b	<i>TN</i> (Doğru negatif) Seçicilik d	b+d
Toplam	a+b	c+d	

Test edilecek değişkenin sürekli olması durumunda belirli bir eşik değer (threshold) kullanılarak değişken iki sonuçlu hale getirilir. Buna göre t gibi bir eşik değer için değişkenin t 'ye eşit ya da t 'den küçük değerleri negatif (D^-), t 'den yüksek olan değerleri ise pozitif (D^+) olarak sınıflandırılır (2).



Şekil 2. Hasta ve Sağlıklı Olasılıkları



Şekil 3. Potansiyel Hasta ve Sağlıklı Olasılıkları

Yapılan sınıflandırmanın doğruluk düzeyinin değerlendirilmesinde Tablo 4’de verilmiş olan doğru atama tablolarından yararlanılır. Tanı testinin doğruluğu hastaları doğru bir şekilde sınıflayan şartlı olasılıklar ile ölçülür.

$X, X_i = \{x_1, x_2, \dots, x_m, \dots, x_n\}$ birim değerlere sahip bir değişken olsun. Sınıflara ve eşik değerlere ilişkin yapılan açıklamalardan hareketle;

$$S^+ \in \{x_1, x_2, \dots, x_m\} \text{ ve } S^- \in \{x_m, x_{m+1}, \dots, x_n\} \quad (3.3)$$

olmak üzere TP, FP, FN ve TN kesirleri belirli bir sınıfa ait kestirilen koşullu olasılıklar ile aşağıdaki gibi ifade edilebilir (2).

Duyarlılık (True Positive Ratio): Bireyin hasta olduğu durumda testin sonucunun pozitif yani hasta olabilme olasılığıdır.

$$TPR = P[D = 1 | S = 1] = P(x_i > t | S^+) = \frac{TP}{TP + FN} \quad (3.4)$$

Seçicilik (True Negative Ratio): Bireyin sağlıklı olduğu durumda testin sonucunun negatif yani sağlıklı olabilme olasılığıdır.

$$TNR = P[D = 0 | S = 0] = P(x_i \leq t | S^-) = \frac{TN}{TN + FP} \quad (3.5)$$

Verilen karar matrisindeki değerlere göre;

$$TNR + FPR = 1 \quad (3.6)$$

$$TPR + FNR = 1 \quad (3.7)$$

eşitliklerinin sağlandığı görülmektedir.

Biyomedikal araştırmalarda test performansını tanımlamak için duyarlılık (TPR) ve seçicilik ($1 - FPR$) sıklıkla kullanılmaktadır (24).

Kesinlik (Accuracy): Tanı testinin doğru sonuç verdiği popülasyon oranıdır.

Hasta popülasyon oranına hastalığın prevalansı (yaygınlığı) denir. Hastalığın prevalansı $\rho = P[S=1]$ şeklinde ifade edilebilir. Ancak kesinlik, tıbbi bir testin tanısal doğruluğunun yeterli bir özeti olarak kabul edilemez. Çünkü hastalığın prevalansından etkilenir (10).

Kesinlik;

$$\begin{aligned} &= TPRx\rho + TNRx(1 - \rho) \\ &= \frac{TP}{TP + FN} x \frac{TP + FN}{TP + TF + FP + FN} + \frac{TN}{FP + TN} x \frac{FP + TN}{TP + TF + FP + FN} \quad (3.8) \\ &= \frac{TP + TN}{TP + TF + FP + FN} \end{aligned}$$

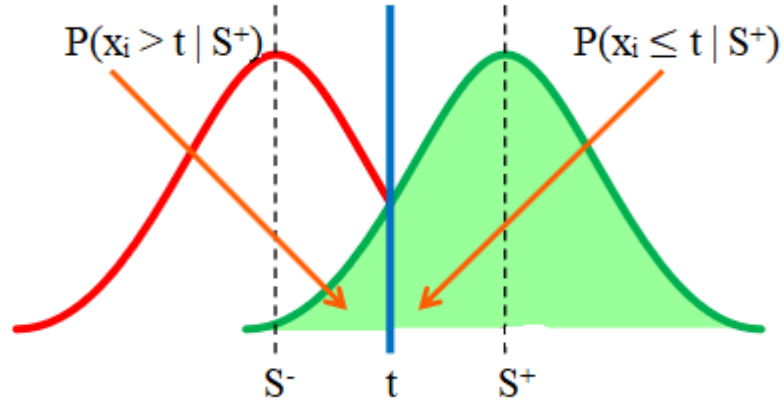
şeklinde ifade edilir.

Yanlış Pozitif Oranı (False Positive Ratio): Bireyin sağlıklı olduğu durumda testin sonucunun pozitif yani hasta olabilme olasılığıdır.

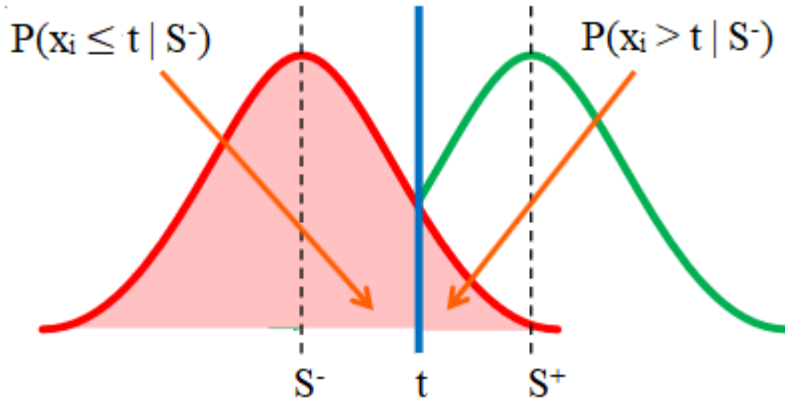
$$FPR = P[D = 1 | S = 0] = P(x_i > t | S^-) = \frac{FP}{TN + FP} \quad (3.9)$$

Yanlış Negatif Oranı (False Negative Ratio): Bireyin hasta olduğu durumda testin sonucunun negatif yani sağlıklı olabilme olasılığıdır.

$$FNR = P[D = 0 | S = 1] = P(x_i \leq t | S^+) = \frac{FN}{TP + FN} \quad (3.10)$$



Şekil 4. Duyarlılık ile Yanlış Negatif Koşullu Olasılıkları



Şekil 5. Seçicilik ile Yanlış Pozitif Koşullu Olasılıkları

Pozitif Tahmin Değeri (Positive Predictive Value): Bireyin tanı testinin sonucunun pozitif olduğu durumda hasta olabilme olasılığıdır.

$$PPV = P[S = 1 | D = 1] = P(S^+ | x_i > t) = \frac{TP}{TP + FP} \quad (3.11)$$

Negatif Tahmin Değeri (Negative Predictive Value): Bireyin tanı testinin sonucunun negatif olduğu durumda sağlıklı olabilme olasılığıdır.

$$NPV = P[S = 0 | D = 0] = P(S^- | x_i \leq t) = \frac{TN}{TN + FN} \quad (3.12)$$

Her hastalık durumu için yanlış sınıflandırma sıklığını göz önünde bulundurmaya alternatif olarak, doğruluk, test sonucunun gerçek hastalık durumunu ne kadar iyi tahmin ettiği ile ölçülebilir.

Tahmin edici (predictive) değerlerin, yalnızca hastalıklı ve sağlıklı bireylerdeki test performansına değil, aynı zamanda hastalık prevalansına bağlı olduğunu görüyoruz. Düşük bir *PPV* düşük hastalık prevalansının bir sonucu olabilir veya gerçek hastalık durumunu çok iyi yansıtamayan bir testten kaynaklanabilir. Tahmini değerler, testin doğal doğruluğunu ölçmek için kullanılmaz. Sınıflandırma olasılıkları, *TPR* ve *FPR*, bu görevle daha ilgili olarak kabul edilir, çünkü testin gerçek hastalık durumunu ne kadar iyi yansıttığını belirlerler. Başka bir deyişle, prediktif değerler testin klinik değerini belirler. Hasta ve klinisyen en çok test sonucu verilen hastalığın mevcut olma ihtimaliyle ilgilenmektedir. Zweign ve Campbell (1993) bu konuda doğruluk ve kullanılabilirlik kavramlarının derinlemesine tartışmışlardır. Birçok araştırma çalışmasında, hem hastalığa özgü sınıflandırma olasılıkları hem de öngörücü değerler bildirilmiştir (9).

Belirttiğimiz bu oran değerleri her ne kadar yanlış pozitif oran ve gerçek pozitif oran olarak adlandırılırsalar da bu ifadeler aslında olasılık veya kesirdir. Oranların özellikle epidemiyoloji araştırmalarında kullanılmasından dolayı ve karışıklığa sebep olmamak için daha uygun olan kesir ifadesi kullanılmıştır

Bir hipotez testi kullanan istatistikçiler, teşhis edici olarak düşünülebilir. Bu durumda kurulan alternatif bir hipoteze ($S=1$) karşı yokluk hipotezinin ($S=0$) istatistiksel hipotez testi bağlamında, anlamlılık düzeyi; *FPR* ve istatistiksel güç; *TPR* terimleri ile ifade edilebilir.

$1-TPR$, hatalı negatif kesre eşit olduğundan (*FPR*, *TPR*) çifti test kullanılırken hataların oluşma ihtimalini tanımlar. İdeal bir testte beklenen kesir değerleri; $FPR=0$, $TPR=1$ şeklindedir. Diğer yandan, işe yaramaz (bir anlam ifade etmeyip sıfır hipotezinin reddedilemediği) bir testte var olan ilişki ise; $TPR=FPR$ şeklindedir.

$1-TPR$ ve *FPR* terimlerinin yanlış sınıflandırma olasılığının her iki bileşeninin doğruluğunu tanımladığı belirtilmelidir. İki tür hatanın maliyet ve sonuçlarının

oldukça farklı olmasıdır. Birincisi, yanlış negatif hatalar yani mevcut olan hastalıklar var iken yanlış tanı sonucu hastalıklar için gerekli tedavinin yapılmamasıdır. Sonuç ölüm kadar ciddi bir durum olabilir. Yanlış pozitif hatalar daha az ciddi olma eğilimindedir. Hastalığı olmayan insanlar yanlış tedavi sonucunda gereksiz tedavi prosedürlerine tabi tutulurlar. Olumsuz etkiler, kişisel rahatsızlık ve stresi içerir.

Ancak uzun vadeli sonuçlar genellikle nispeten küçüktür. Sınıflandırma olasılık çiftini (*TPR*, *FPR*) tanısal doğruluğu özetlemek için genel yanlış sınıflandırma olasılığına tercih edilmesinin diğer bir sebebi ise, yanlış sınıflandırma olasılığının hastalık prevalansına bağlı olmasıdır. Çünkü yanlış sınıflandırmadan dolayı bilgilendirici olmayan bir testin uygulandığı popülasyon prevalansı düşük ise bu test düşük bir genel hata olasılığına sahip olacaktır (2, 23).

2.2.2. Olabilirlik Oranı (Likelihood Ratio)

Tanı testi sonuçları hastalık hakkında karar vermek üzere kullanılırken dikkat edilmesi gereken kısım, sadece duyarlılık ve seçicilik ölçütlerinin ele alınmamasıdır. Bu ölçütler ön olasılıklar ile değerlendirildiğinde anlam taşımaktadır. Klinisyen tarafından en iyi tahmin olabilirlik oranlarını, duyarlılık ve seçicilik ile birlikte değerlendirmek ile yapılabilir.

Olabilirlik oranları, bir testin prognostik veya tanısal değerini tanımlamanın başka bir yoludur. Olabilirlik oranının hesaplamasında ön olasılıkların yerine ön görelilik (odds ratio) değerler kullanılır. Hastalığın test sonrası olasılığını hesaplamının bir diğer yolu da odds-olabilirlik hesaplama yaklaşımıdır. Burada ele alınan konu, test öncesi olasılığı ile testin duyarlılık ve seçiciliğinin hastalığın test sonrası olasılığı üzerine etkisini belirlemektir. Hastalığın test öncesi olasılığı ve hastalığın prevalansı, test sonrası olasılığa önemli etkide bulunur. Sackett ve ark. (1985), Boyko (1994) ve Boyd (1997), tanısal test değerlendirmesi için önemli bir ölçüt olduğunu savunmuşlar, çalışmalarında bu konu üzerine tartışmışlardır (23, 25, 26).

Bu parametreler, gerçek istatistiksel anlamdaki olasılık oranlarıdır. Buna rağmen bağlamı ayırt etmek için tanısal olasılık oranları olarak adlandırıyoruz. Tanı

testi sonuçları iki olasılıklı olmasından dolayı pozitif teste ve negatif teste ait iki tane olabilirlik oranı vardır.

Pozitif Olabilirlik Oranı (Positive Likelihood Ratio): Tanı testi sonucunun hasta olan kişide pozitif çıkma olasılığının sağlıklı kişide pozitif çıkma olasılığına oranıdır.

$$LR^+ = \frac{P[D = 1 | S = 1]}{P[D = 1 | S = 0]} = \frac{P(D^+ | S^+)}{P(D^+ | S^-)} = \frac{TPR}{1 - TNR} = \frac{Duyarlılık}{1 - Seçicilik} \quad (3.13)$$

Negatif Olabilirlik Oranı (Negative Likelihood Ratio): Tanı testi sonucunun hasta olan kişide negatif çıkma olasılığının sağlıklı kişide negatif çıkma olasılığına oranıdır.

$$LR^- = \frac{P[D = 0 | S = 1]}{P[D = 0 | S = 0]} = \frac{P(D^- | S^+)}{P(D^- | S^-)} = \frac{TNR}{1 - TPR} = \frac{1 - Duyarlılık}{Seçicilik} \quad (3.14)$$

Tanı testlerinin ideal bir test olması için pozitif olabilirlik oranının yüksek, negatif olabilirlik oranının ise düşük olması istenir.

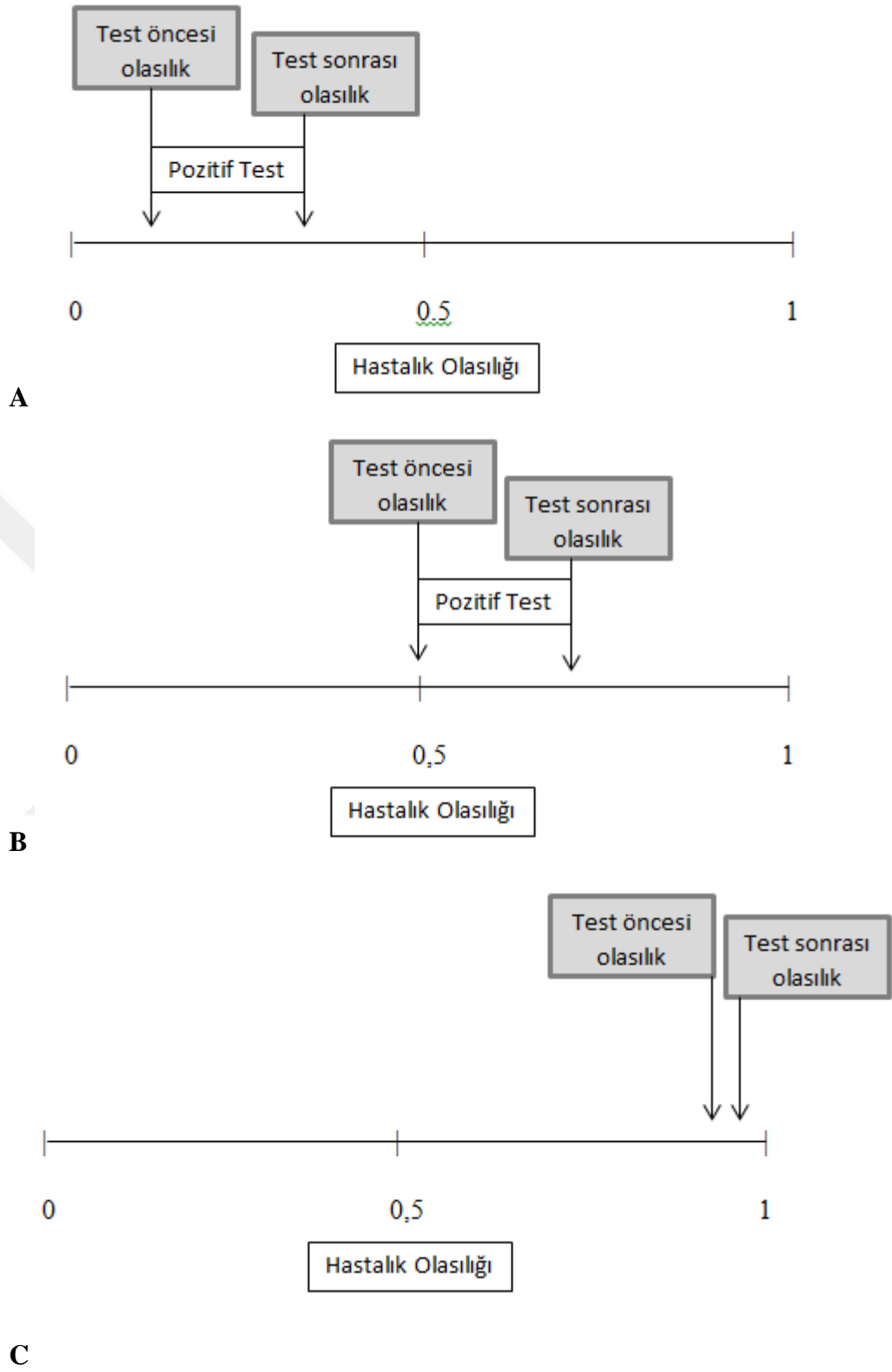
Olabilirlik oran metodunun en önemli kullanım avantajı, sadece bir değeri göz önünde bulundurmasıdır. Duyarlılık ve seçicilik gibi değerlerinin hesaplanmasına göre daha kolay anlaşılmakta ve prevalansdaki değişimlere göre duyarlılık ve seçicilikten daha güçlü sonuçlar vermektedir (16).

Aşağıda verilmiş olan Şekil 6'da A, B ve C şekillerini incelediğimizde;

A şeklinde; uygulanan tanı testi ile ilgili verilen bilgiler neticesinde, tanının mümkün olmadığı yorumu çıkarılabilir.

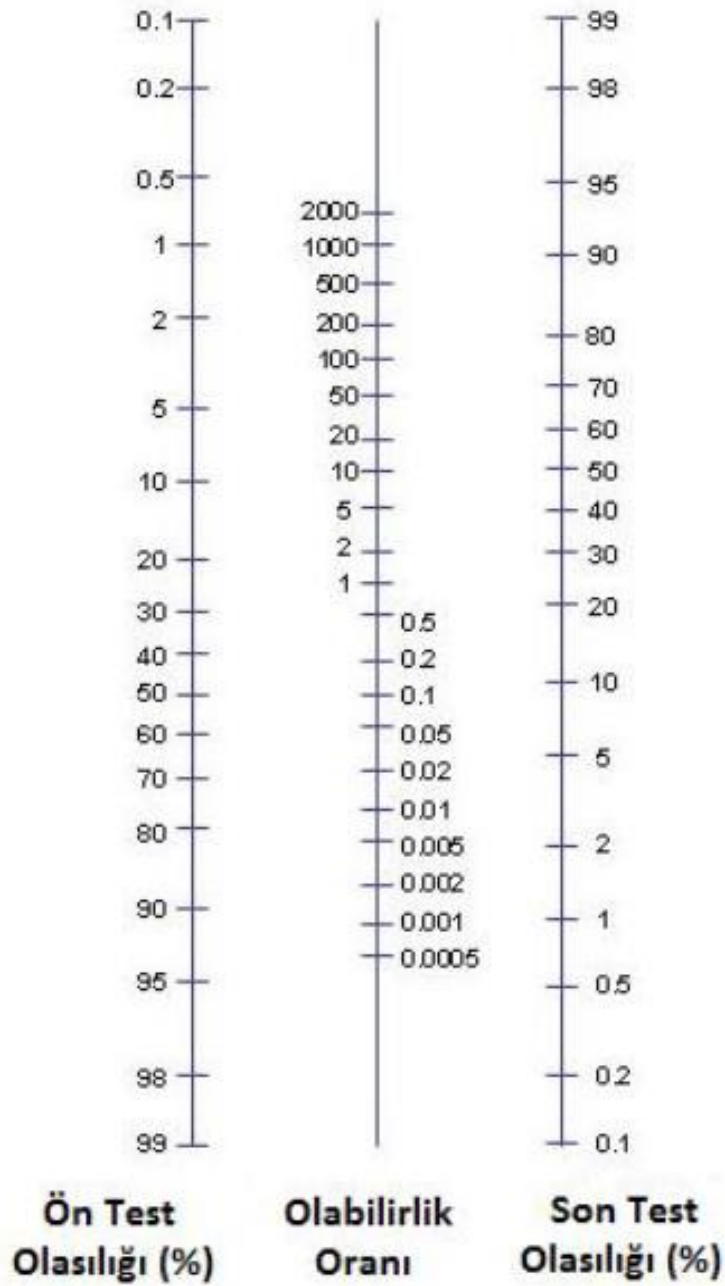
B şeklinde; uygulanan tanı testi ile ilgili verilen bilgiler neticesinde tanının belirsiz olduğu yorumu çıkarılabilir.

C şeklinde; uygulanan tanı testi ile ilgili verilen bilgiler neticesinde tanının kesin olduğu yorumu çıkarılabilir.



Şekil 6. Test Olasılıkları ile Hastalık Olasılığı Arasındaki İlişki

Test öncesi olasılıklar ve olabilirlik oranları kullanılarak test sonrası olasılıkların hesabı için Fagan (1975) tarafından yapılmış Fagan Nomogramı kullanılabilir. Nomogram üzerinde 3 adet skala bulunmaktadır; sırasıyla test öncesi olasılık, olabilirlik oranı ve test sonrası olasılıklar. Test öncesi olasılık değeri ve olabilirlik oranı üzerinden geçecek bir çizgi çizilerek olası test sonrası olasılığı bulunabilir (27).



Şekil 7. Fagan Nomogramı

Şimdi ele aldığımız ölçütleri bir tabloda özetleyelim;

Tablo 5. Tanı Testi Değerlendirme Ölçütleri

	Sınıflandırma Olasılıkları	Tahmini Değerler	Tanı Olabilirlik Oranları
Değişkenler	$FPR=P[D=1/S=0]$ $TPR=P[D=1/S=1]$	$PPV=P[S=1/D=1]$ $NPV=P[S=0/D=0]$	$DLR^+=\mathcal{L}R(D=1)^*$ $DLR^-=\mathcal{L}R(D=0)^*$
Değer Aralığı	(0,1)	(0,1)	(0, ∞)
Mükemmel Test	$FPR=0$, $TPR=1$	$PPV=1$, $NPV=1$	$DLR^+=\infty$, $DLR^-=0$
Kullanışsız Test	$FPR=TPR$	$PPV=\rho$, $NPV=1-\rho$	$DLR^+=1=DLR^-$
Kullanım İçeriği	Doğruluk	Klinik Tahmin	Bilgilendirme Testi
Cevap Verdiği Soru	Test gerçek hastalık durumunu ne dereceye kadar yansıtır?	Test sonucu verilen hastalık ne kadar muhtemel?	Test hastalık durumu bilgisini ne kadar değiştiriyor?
Hastalık Prevalansından Etkilenir mi?	Hayır	Evet	Hayır

$$*\mathcal{L}R(y) = P[Y=y | S=1]/P[Y=y | S=0]$$

2.2.3. Bayes Teoremi

Tanı testinin kestirim (tahmin) değerinin hesaplanmasında son metot olarak Bayes Teoremi incelenecektir. Matematiksel bir metot olan Bayes Teoremi, 18. YY'da matematikçi Thomas Bayes tarafından geliştirilmiştir. Bayes Teoremi epidemiyoloji biliminin ilk yıllarında olasılıkların yeniden hesaplanmasında oldukça sık yararlanılan bir metottur (23).

Ele aldığımız 2x2 Atama Tablosu ve Olabilirlik Oranlarını temel alan metotlarımız Bayes Teoremine dayanan metoda göre daha sık kullanılan yöntemlerdir.

Bayes Teoremini kullanarak yapılan hesaplama ile pozitif teste ait kestirim değeri bulunur. Yani bireye uygulanan tanı testi sonucunun pozitif olduğunda

gerçekte hasta olma olasılığı hesaplanır. Hesaplanan olasılık diğer metotlarda olduğu gibi koşullu olasılık hesabı ile yapılmaktadır. Bayes Teoremi kullanılarak oluşturulan formül;

$$P(S^+ | D^+) = \frac{P(D^+ | S^+)P(S^+)}{P(D^+ | S^+)P(S^+) + P(D^+ | S^-)P(S^-)} \quad (3.15)$$

$$P(S^- | D^-) = \frac{P(D^- | S^-)P(S^-)}{P(D^- | S^-)P(S^-) + P(D^- | S^+)P(S^+)} \quad (3.16)$$

şeklinindedir. Buradaki koşullu olasılıklar, $P(D^+ / S^+) =$ Duyarlılık, $P(D^- / S^-) =$ Seçicilik $P(S^+) =$ Ön olasılık, $P(D^+ / S^-) =$ Yanlış pozitif oranı $P(D^- / S^+) =$ Yanlış negatif oranı, eşitliklerini ifade etmektedir.

2.3. ROC Analizi

ROC analizi, bir tanı testi olarak, bir işaretleyici veya bir sınıflandırıcıyı temsil eden sürekli bir değişkenin ayırt edici performansını değerlendirmek için yaygın olarak kullanılır. Bu bölümde ele aldığımız analizin ana unsuru olan ROC eğrisinin; özellikleri, eğrinin yorumlanmasında yararlanılan göstergeler, bu göstergelerin farklı yaklaşımları ve de kullanım alanlarından bahsedilecektir.

2.3.1. ROC Eğrisi

ROC eğrisi, bir tanı testinin yanlış pozitif ve gerçek pozitif oranlar arasındaki ilişkinin grafiksel bir gösterimidir. Yani, duyarlılık ve 1-seçicilik ile oluşturulmuş bir fonksiyonun grafiksel şeklidir. Duyarlılık, gerçekten hastalıklı bir bireyin pozitif bir test sonucuna sahip olma olasılığı ve seçicilik ise, gerçekten hastalıklı bir bireyin, negatif bir test sonucuna sahip olma olasılığıdır (2).

ROC eğrisi, çeşitli eşik (threshold) ayarlarında sınıflandırma problemi için bir performans ölçümüdür. ROC bir olasılık eğrisidir. Bazı temel özellikleri incelemek için, ROC eğrisini daha önceki bölümlerde elde aldığımız oran ve eşik değerleri ile

tanımlayalım. ROC eğrisi koordinat ekseninin birinci bölgesinde yer alan monoton artan bir fonksiyondur. Bir t eşiği kullanarak, sürekli test sonucu (x_i) ile ilişkisine bağlı olarak;

$$“x_i \geq t \text{ ise pozitif}” \quad (3.17)$$

$$“x_i < t \text{ ise negatif}” \quad (3.18)$$

sınıflanan; ikili bir test tanımlayalım. t eşik değerinde karşılık gelen doğru ve yanlış pozitif oranlar sırasıyla $TPR(t)$ ve $FPR(t)$ olsun. Buna göre eşik değerler ile kestirilen olasılıklar arasında aşağıdaki gibi bir bağıntı kurulabilir.

$$\begin{aligned} TPR(t) &= P(x_i \geq t | S = 1) \\ FPR(t) &= P(x_i \geq t | S = 0) \end{aligned} \quad (3.19)$$

ROC eğrisi, x 'i farklı eşiklerde dikotomize ederek elde edilebilecek muhtemel doğru ve yanlış pozitif kesirlerin tamamıdır. Yani ROC eğrisi;

$$ROC(.) = \{(FPR(t), TPR(t)), t \in (-\infty, +\infty)\} \text{ 'dir.} \quad (3.20)$$

Eşik değeri arttıkça hem TPF(t) hem de FPF(t) değerlerinin azaldığını görülür. En uç değeri alan t eşik değerinde;

$$\begin{aligned} \lim_{t \rightarrow \infty} TPR(t) &= 0 & \lim_{t \rightarrow -\infty} TPR(t) &= 1 \\ \lim_{t \rightarrow \infty} FPR(t) &= 0 & \lim_{t \rightarrow -\infty} FPR(t) &= 1 \end{aligned} \quad (3.21)$$

eşitlikleri sağlanır. Bu yüzden ROC eğrisi koordinat sisteminde birinci bölgede yer alan monoton artan bir eğri fonksiyonudur. t , $FPR(t) = c$ 'nin eşik değeri ve $TPR(t)$ 'nin ROC fonksiyonundaki karşılığı olmak üzere;

$$ROC(.) = \{(c, ROC(c)), c \in (0, 1)\} \quad (3.22)$$

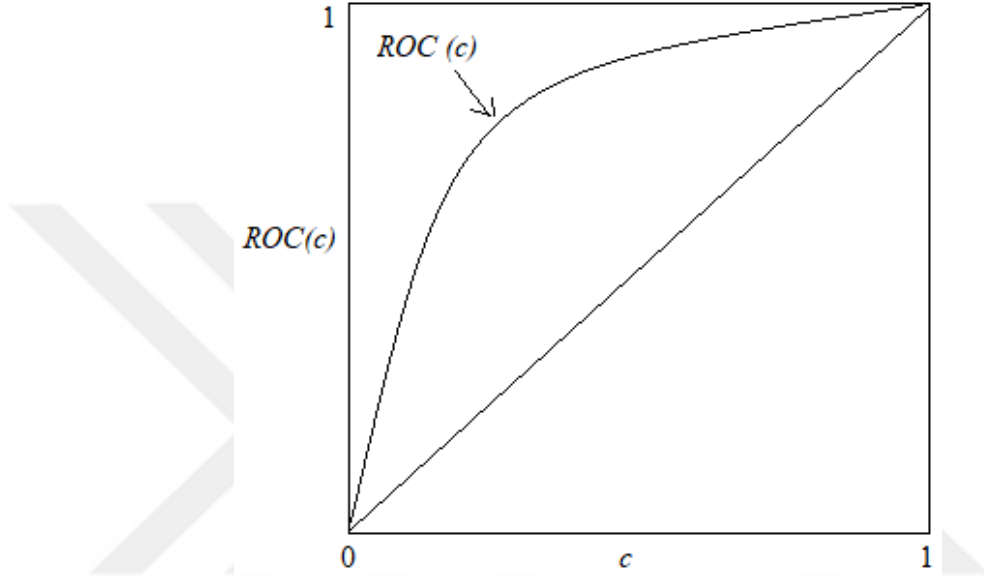
eşitliği mümkündür.

Bilgilendirici olmayan bir tanı testinde ise, olasılık dağılımları hasta ve sağlıklı popülasyonlarda aynıdır. Bu nedenle herhangi bir t eşik değeri için

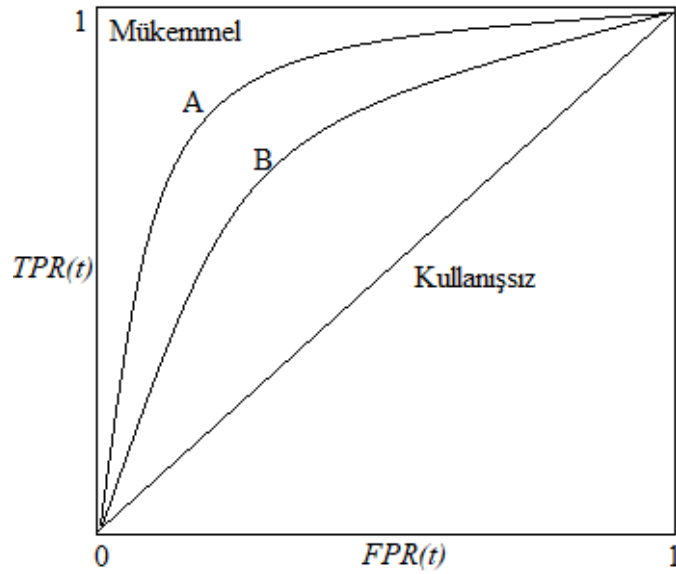
$$TPR(t) = FPR(t)'dir. \quad (3.23)$$

Bilgilendirici olmayan bir tanı testinin ROC eğrisi de birim eğime sahip $ROC(t)=t$ eşitliğine sahiptir. Mükemmel bir sonuç veren tanı testi ise hasta ve sağlıklı bireyleri tamamen ayırmaktadır. Bazı t eşik değerleri için

$$TPR(t) = 1, FPR(t) = 0'dır. \quad (3.23)$$

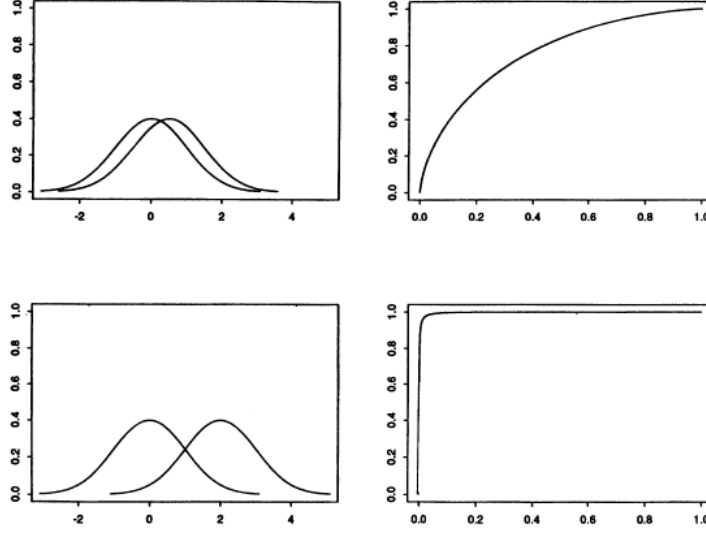


Şekil 8. ROC Eğrisi Örneği



Şekil 9. İki Test için ROC Eğrileri

Aşağıda dağılımları verilmiş tanı testi ile altın standart testlerinin sonuçlarının kestirim güçlerinin artması neticesinde değişen duyarlılık ve seçicilik değerlerine bağlı olarak ROC eğrisinin etkinlikleri gösterilmiştir.



Şekil 10. Test Sonuçlarının Kestirim Gücü ile ROC Eğrisi Arasındaki İlişki

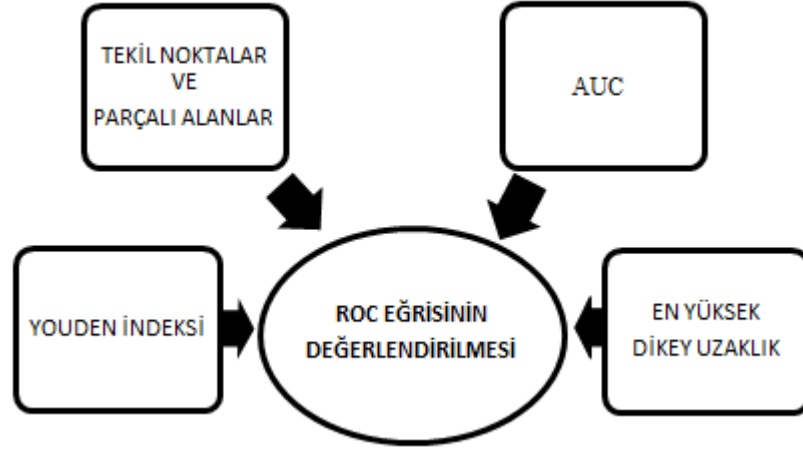
Tanı testlerini değerlendirmek için kullanılan ROC eğrisinin bazı özellikleri aşağıda verilmiştir.

1. Potansiyel performansın tam bir tanımını sunar.
2. Aynı testin çalışmaları arasında bilgilerin karşılaştırılmasını ve birleştirilmesini kolaylaştırır.
3. Uygulamalarda eşik seçimini yönlendirir.
4. Farklı non-binary testler arasında yapılan karşılaştırmalar için bir mekanizma sağlar (23).

2.3.2. ROC Eğrisinin Değerlendirilmesinde Kullanılan Göstergeler

ROC eğrileri için sayısal indeksler genellikle eğrileri özetlemek için kullanılır ve bu tür indeksler eğri hakkında önemli bilgiler verir. Bu, olasılık dağılımını tanımlamak için ortalama ve varyans gibi tanımlayıcı istatistiklerin kullanılmasına

benzer. Özet indeksleri, özellikle pek çok test yapıldığında yararlıdır. Ayrıca, ROC eğrilerinin karşılaştırılmasında çıkarımsal istatistiklerin temeli olarak kullanılabilirler.



Şekil 11. ROC Analizinde Değerlendirme Göstergeleri

ROC eğrilerinin değerlendirilmesinde sıklıkla kullanılan ve en bilinen indeksler AUC, Tekil Noktalar ve Parçalı Alanlar, Youden İndeksi (YI) ve En Yüksek Dikey Uzaklık (MVD)'dir (10).

2.3.2.1. Eğri Altında Kalan Alan (AUC)

AUC, ROC eğrisinin performans değerlendirmesinde ve farklı analizler neticesinde elde edilen kestirimlerin doğruluklarının ya da performanslarının kıyaslanması için en yaygın kullanılan göstergedir. AUC,

$$AUC = \int_0^1 ROC(c)dc \quad (3.24)$$

şeklinde tanımlanır. AUC indeksinin tanım aralığı,

$$0,5 \leq AUC \leq 1 \text{ 'dir.} \quad (3.25)$$

ROC eğrisinin bulunduğu koordinat bölgesinin yapısı temel alınacak olursa, eksenleri 0'dan 1'e kadar ölçümlenmiş bir kare içerisine çizilen eğrinin alan

hesabından yola çıkılmak üzere üst sınır 1 alt sınır ise 0'dır. Pozitif ve Negatif ayrımı en mükemmel şekilde yapmak için bu sınırları oluşturur. AUC değeri 1 olduğunda en iyi ayırım gücüne, AUC değeri 0,5 olduğunda ise en kötü ayırım gücüne sahiptir. Yani bir anlamlılığı olmadığı, sıfır hipotezinin geçerli olduğu şeklinde ifade edilebilir. Tanı testlerinin mükemmel olduğu sonucuna bulunan AUC değerinin 1 ya da 1'e yakın olduğu durumlarda varılır. Bilgilendirici olmayan bir tanı testinde ise AUC değeri 0,5'dir (20).

Tablo 6. AUC Değerlerinin Yorumlanması

AUC = 0,5	Tanı testi ayırım gücüne sahip değildir.
0,5 < AUC < 0,7	Tanı testi "zayıf" bir ayırım gücüne sahiptir.
0,7 ≤ AUC ≤ 0,8	Tanı testi "kabul edilebilir" bir ayırım gücüne sahiptir.
0,8 ≤ AUC ≤ 0,9	Tanı testi "mükemmel" bir ayırım gücüne sahiptir.
0,9 ≤ AUC	Tanı testi "olağanüstü" bir ayırım gücüne sahiptir.

2.3.2.1.1. Parametrik Yaklaşımına Göre AUC Kestirimi

ROC eğrisinin meydana gelmesinde kullanılan koşullu olasılıklar parametrik yaklaşımlar göz önünde bulundurularak elde edildiği söylenebilir. Bu yaklaşıma göre sürekli değişkenler için elde edilen koşullu olasılık kestirimlerinin iki değişkenli normallik varsayımını sağladığı düşünülür.

İkili normal dağılıma sahip ROC eğrisi koşullu olasılıklar kullanılarak alan hesaplamalarıyla çizilebilmesinin yanı sıra bir fonksiyon yardımıyla da çizilebilir. Hasta ve sağlıklı gruplardaki sürekli test sonuçları normal dağılım gösterdiği ele alınırsa, ROC eğrisi fonksiyonu;

$$\left(X_{D^+} \sim (\mu_{D^+}, \sigma_{D^+}), X_{D^-} \sim (\mu_{D^-}, \sigma_{D^-}) \right) \quad (3.26)$$

parametreler ile bu şekilde ifade edilebilir. ROC eğrisi fonksiyonu,

$$ROC(c) = \phi(a + b\phi^{-1}(c)) \quad (3.27)$$

şeklinde tanımlanabilir. Yukarıdaki eşitlikte “ c ” FPR , “ a ” sabit, “ b ” eğrinin eğimini ifade eder. TPR oranını gösteren $ROC(c)$ fonksiyonun kestirimi için;

$$a = \frac{(\mu_{D^+} - \mu_{D^-})}{\sigma_{D^+}} \quad (3.28)$$

$$b = \frac{\sigma_{D^-}}{\sigma_{D^+}}$$

eşitliklerinden yararlanılır. ϕ , birikimli standart normal dağılım fonksiyonudur. t gibi bir eşik değerinde ROC eğrisinin koordinatlarını oluşturan;

$$TPR(t) = P(X_{D^+} > t) = \phi\left(\frac{\mu_{D^+} - t}{\sigma_{D^+}}\right)$$

$$FPR(t) = P(X_{D^-} > t) = \phi\left(\frac{\mu_{D^-} - t}{\sigma_{D^-}}\right) \quad (3.29)$$

şeklinindedir. FPF , “ c ” ile gösterilmek üzere; testin eşik değeri,

$$t = \mu_{D^-} - \sigma_{D^-} \phi^{-1}(c) \text{ 'dir.} \quad (3.30)$$

Sabit ve eğimi belirlendikten sonra;

$$ROC(c) = TPR(t) = \phi\left(\frac{\mu_{D^+} - t}{\sigma_{D^+}}\right)$$

$$ROC(c) = \phi\left(\frac{\mu_{D^+} - \mu_{D^-} + \sigma_{D^-} \phi^{-1}(c)}{\sigma_{D^+}}\right)$$

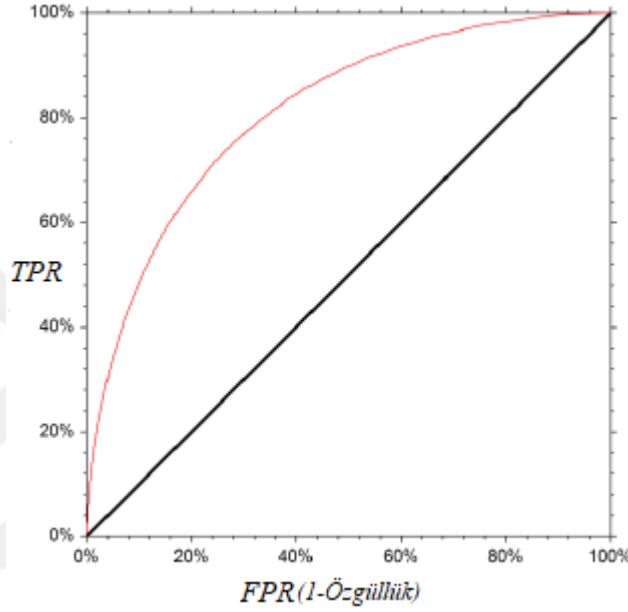
$$ROC(c) = \phi(a + b\phi^{-1}(c)) \quad (3.31)$$

elde edilir. AUC indeksi, ROC eğrisi binormal olduğunda basit bir analitik forma sahiptir. Yukarıdaki bilgilere dayanarak AUC;

$$AUC = \phi\left(\frac{a}{\sqrt{1+b^2}}\right) \quad (3.32)$$

eşitliğine sahiptir (10).

Aşağıda parametrik olan yaklaşıma göre kestirilmiş bir ROC eğrisi gösterilmiştir.



Şekil 12. Parametrik Olan Yaklaşıma Göre Kestirilmiş ROC Eğrisi

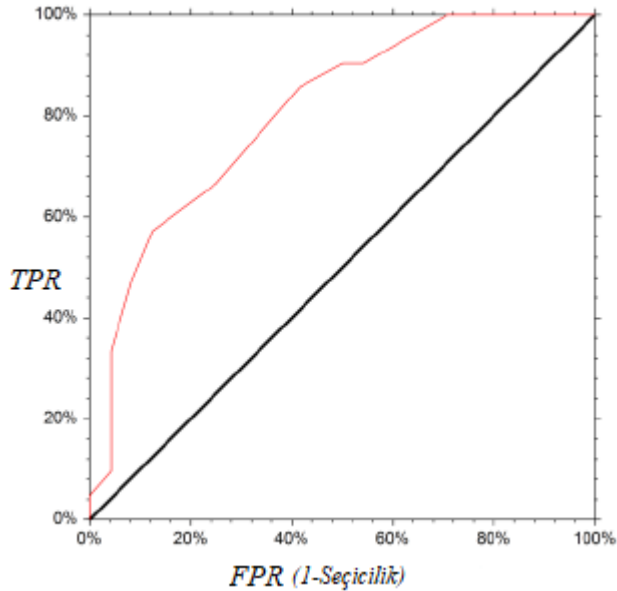
2.3.2.1.2. Parametrik Olmayan Yaklaşıma Göre AUC Kestirimi

Normal dağılım göstermeyen veriler için ROC eğrisi kestiriminin parametrik yaklaşıma göre yapılması hatalı sonuçların ortaya çıkmasına neden olmaktadır. Bu tür veriler için yapılacak ROC eğrisi kestiriminde herhangi bir dağılım varsayımı içermeyen parametrik olmayan yaklaşımlar geliştirilmiştir. Ampirik yöntem ve düzgünleştirilmiş eğri yöntemleri bunlardan birkaçıdır. Parametrik olmayan yaklaşım parametrik yaklaşıma göre esnek bir yapıya sahiptir. Bu nedenle de bu yaklaşımda AUC indeksini hesaplamak daha kolaydır. Bu yaklaşımın dezavantajı ROC eğrisinin düzgünleştirilmiş kestiriminin olmamasıdır. Yani eğri kırılmalar meydana gelerek oluşmuştur. Bu yüzden de bilgi kaybına sebep olabilmektedir.

Ampirik yöntemde, ROC eğrisini düz çizgilerin veri noktalarını (duyarlılık, 1-seçicilik) bağlayarak ve sonra trapez kuralını kullanarak tahmini alan AUC hesaplanmaktadır. Bu şekilde hesaplanan tahmini AUC'nin, hasta ve sağlıklı değerlerin olası çiftleri sayısı ile normalize edilen Mann-Whitney U bağımsız test istatistiklerine eşdeğer olduğu gösterilmiştir. Ayrıca iki örnekli durum Wilcoxon Rank-Sum test sıra toplamı istatistiği olarak da bilinmektedir (10).

Düzleştirilmiş eğri yönteminde, gerçek bir ROC eğrisi düzgündür, ancak ampirik yöntem ROC noktalarını düz çizgi parçalarıyla bağladığından, tahmin edilen eğri pürüzlüdür. Bu nedenle, veriye herhangi bir parametrik varsayım getirmeden ampirik eğriyi basitleştirerek gerçek eğri tahminimizi geliştirebiliriz. Bu yaklaşım, her veri noktasının konumuna yatay eksen boyunca bir çekirdek yoğunluğu işlevi yerleştirerek ve işlevleri toplayarak bir histogramı pürüzsüz hale getirmektedir. Çekirdeğin yatay ölçeklendirilmesi, düzleştirme derecesini kontrol eder ve bir bant genişliği parametresi h ile belirtilir. Her bir çekirdek fonksiyonunun altındaki alan $1/n$ 'dir, bu nedenle toplanan alan her zaman bütündür (2, 23).

Aşağıda parametrik olmayan yaklaşıma göre kestirilmiş bir ROC eğrisi gösterilmiştir.



Şekil 13. Parametrik Olmayan Yaklaşıma Göre Kestirilmiş ROC Eğrisi

2.3.2.1.3. Yarı Parametrik Olan Yaklaşım Göre AUC Kestirimi

Yarı parametrik yöntem, parametrik ve parametrik olmayan yöntemlere göre ROC eğrisi altında kalan alan hesabında, ortaya çıkan dezavantajlardan kurtulmak için alternatif olarak geliştirilmiş yaklaşımdır. ROC eğrisi altında kalan alan kestirimi, yarı parametrik olan yaklaşıma göre yapıldığında elde edilen sonuç parametrik ve parametrik olmayan yaklaşıma göre yapıldığında elde edilen sonuçtan daha güçlüdür. Bundan dolayı yarı parametrik olan yaklaşıma göre kestirilen ROC eğrisi parametrik olmayan yaklaşıma göre kestirilen eğriden daha düzgündür.

ROC eğrisi kestirimi için kullanılan yarı parametrik yaklaşım üzerine farklı araştırmacılar tarafından çalışmalar yapılmıştır. Bu çalışmaların bazılarında araştırmacılar tarafından geliştirilen çeşitli yaklaşımlar mevcuttur. Green ve Swets tarafından geliştirilen, belirli monoton artış dönüşümlerinin sonunda test değişkeninin iki değişkenli normal dağılacığı varsayımı üzerine kurulmuş iki değişkenli normal model yaklaşımıdır (28).

ROC analizi için bir dizi yazılım paketi mevcuttur. R, SAS, SPSS, MedCalc, AccuROC, NCSS, ROCKIT, Analyse-It, CMDT, GraphROC, LABROC bunlardan bazılarıdır. Bazı araştırmacılar bu yazılımları doğruluk, eksiksizlik ve kullanım kolaylığı açısından değerlendirmişlerdir. Tek bir paketin tüm ihtiyaçlar için yeterli olmadığını bulmuş, ancak varsayımlarını ve yöntemlerini dikkate alarak tüm doğru hesaplamaları yapmıştır. En son yazılımlardan olan ROCKIT, yıllardır ücretsiz ROC analiz yazılımını istatistiksel topluluğa katkıda bulunan bir araştırma laboratuvarının ürünüdür (29).

Birçok kestirim için kullanılan en bilinen yazılım olan SPSS paket programında ise, sadece parametrik olmayan yaklaşım mevcut olup Mann Whitney-U yöntemi kullanıldıktan ROC eğrisi kestirimi yapılmaktadır. Parametrik olan yaklaşım ve yarı parametrik olan yaklaşımların mevcut olduğu, yukarıda bahsettiğimiz LABROC yazılımı ROC eğrilerinin kestirimi için kullanılabilir.

2.3.2.2. Tekil Noktalar ve Parçalı Alanlar

Belirli bir yanlış pozitif kesri (FPR), $ROC(c_0)$ yani TPR ile ilgili bir özet endeksi sağlar. Bazı mevcut kaynaklar için yanlış pozitif bir oranını barındırmak mümkün olabilir. Eşikleri, her biri $FPR = c_0$ 'ı verecek şekilde olduğu zaman testleri karşılaştırmak, $ROC(c_0)$ değerlerini karşılaştırarak gerçekleştirilir. ROC eğrisi altındaki kısmi alan;

$$PAUC(0, c_0) = \int_0^{c_0} ROC(c)dc \quad (3.33)$$

şeklinde hesaplanır (20).

2.3.2.3. Youden İndeksi (YI)

Youden indeksi (YI), AUC ile tekil noktalar ve parçalı alanlar kadar çok yaygın kullanılsa da diğerlerinden sonra en yaygın kullanılan indekslerden birisidir. YI sürekli değişkenlerde TPR ile FPR ölçütleri arasındaki maksimum farktır (20, 22).

$$YI = \max (TPR - FPR) \quad (3.34)$$

eşitliğine sahiptir. $FPR + TNR = 1$ olduğundan dolayı;

$$YI = \max (TPR + TNR - 1) \quad (3.35)$$

$$YI = \max (\%Duyarlılık + (\%Özgüllük-100)) \quad (3.36)$$

denkliğine ulaşılır.

TPR ile FPR ölçütleri arasındaki farkın maksimum olduğu nokta YI optimal ayırım gücüne sahiptir. c eşik değeri de en iyi ayırım gücünü veren değerdir (23).

2.3.2.4. En Yüksek Dikey Uzaklık (MVD)

En yüksek dikey uzaklık (MVD), ROC eğrisi ile 45° çizgisi arasındaki maksimum dikey uzaklığı veren indekstir. ROC eğrisinin çizildiği koordinat bölgesi göz önüne alınacak olursa; indeksin tanım aralığı [0, 1]'dir. Bu indeks Kolmogorov-Smirnov'un iki dağılım arasındaki mesafeyi ölçmesi ile ilişkilidir (30).

En yüksek dikey uzaklık;

$$MVD = \max_c |ROC(c) - c| \quad (3.37)$$

$$MVD = |TPR - FPR| \quad (3.38)$$

şeklinde ifade edilir (9).

2.3.3. ROC Eğrisinin Kullanım Alanları

ROC eğrilerinin özelliklerinde bahsettiğimiz üzere, eğrinin oluşumunda etkili olan duyarlılık ve seçicilik ölçütlerinden yola çıkarak eğrinin en iyi kestirimi yapması sağlanmaktadır. Birincil olarak optimal kestirimi elde etmek için optimal duyarlılık ve seçicilik ölçütlerinin belirlenmesi için bir değere ihtiyaç duyulduğu görülmektedir. İkincil olarak ROC eğrisi kestiriminde kullanılan göstergelerin ise eğrinin kestirim gücünü gösterdiği görülmektedir. Bunlardan yola çıkarak ROC eğrisi, optimal duyarlılık ve seçiciliği belirleyen eşik değeri belirlemede ve kestirim gücünün göstergelerini kullanarak farklı testler arasında performans kıyaslamalarında kullanılabileceği çıkarımı yapılabilir.

2.3.3.1. Optimum Eşik Değeri (Cutoff Value) Belirleme

ROC eğrisi üzerinde bulunacak optimum kesim noktası tanı testinin en doğru sınıflamayı yapmasını sağlayan noktadır. Bu noktayı bulmak için bilinen iki farklı kriter kullanılır. Bunlar eğri üzerindeki (0,1) noktasına en yakın nokta ve de YI kriterleridir. Bu iki kriterde duyarlılık ve seçiciliğe aynı derecede önem verilir.

Optimum eşik değerini belirlemek için, ROC eğrisi üzerinde olan herhangi bir nokta ile (0,1) noktası arasındaki uzaklık eğri üzerindeki gözlenen tüm kesim noktaları için hesaplanır. Bu noktalardan (0,1) noktasına en yakın olanı optimum eşik değerini vermektedir (10).

YI eğri üzerindeki bir noktanın (0,0) ve (1,1) doğrularını birleştiren ana köşegene uzaklığıyla ilgilidir. Bu uzaklığın maksimum değeri YI verir. Farklı duyarlılık ve seçicilik değerleri için farklı YI bulunabilir. YI doğru sınıflama oranını maksimum olma eğilimini gösterdiği için tercih edilen bir kriterdir.

2.3.3.2. Farklı Tanı Testlerinin Performanslarının Değerlendirilmesi

ROC eğrilerinin kestirim gücünü belirtmek için farklı indislerin ya da göstergelerin olduğundan bahsedilmişti. Bu göstergeleri kullanarak farklı testler arasında performans kıyaslamalarında kullanılabileceği çıkarımı yapılabilir.

Belirli bir vaka üzerinde tanı koymada kullanılan testlerin çeşitliliği fazla olabilir. Bu tanı testlerinin, hasta ile sağlıklı ayrımını en doğru sınıflama ile yapan testi tercih edilir. Bu seçimi yapabilmek adına var olan testlerin performans değerlendirilmeleri ise belirlenen göstergenin kıyaslanması ile olmaktadır. ROC eğrilerinde sıklıkla kullanılan AUC göstergesi temel alındığında, bu gösterge değerlerinin kıyaslanması için var olan eğriler altında kalan alanlar arasında farklılığın olup olmaması araştırılmalıdır. Bunun içinde konunun başında da belirttiğimiz gibi z dağılımından yararlanılır (22, 23).

AUC göstergelerinin karşılaştırmalarını hipotez kontrolü üzerinden değerlendirecek olursak;

$$\begin{aligned} H_0 : AUC_1 - AUC_2 &= 0 \\ H_1 : AUC_1 - AUC_2 &\neq 0 \end{aligned} \quad (3.39)$$

şeklinde kurulur. Kestirimleri karşılaştırılacak olan eğriler aynı bireyler üzerinde farklı tanı testleri olarak yer almakta ise her iki tanı testinin oluşturduğu gruplar

arasında bir bağımlılık meydana gelir. Yapılacak olan hipotez kontrolünde kullanılacak olan z değeri aşağıdaki gibi hesaplanır.

$$z = \frac{AUC_1 - AUC_2}{\sqrt{\text{var}(AUC_1) + \text{var}(AUC_2) - 2\text{Cov}(AUC_1, AUC_2)}} \quad (3.40)$$

Kullanılan yazılımlar neticesinde var olan z değerine göre bulunan p değeri, tanı testlerini temsil eden eğrilerin kestirim güçlerini yani testlerin performansları arasında farklılığın olup olmadığını tayin eder. Yorumlamalar yapılırken denklemden de görüleceği üzere hesaplanan standart hatalar önemlidir ve bu hatalar dahil edilmelidir.



3. GEREÇ VE YÖNTEM

Bu bölümde araştırmanın modeli, kullanılacak çalışma grubu, verilerin analizi ve verilen toplanması üzerine genel bilgiler verilmiştir.

3.1. Araştırmanın Modeli

Bu çalışmada tanı değerlendirme ölçütleri ile ROC analizini tanıtılması ve anlatılan, verilen bilgilerin klinik vakalar üzerinden farklı istatistiksel programlar ile uygulanması ve çalışmada yer verilen diyabet ve akut pankreatit vakalarında kullanılan tanı testlerinin performanslarını ve hastalık durumlarının en yüksek doğru sınıflama oranları ile ayırt edebileceği eşik değerlerini (kesim noktalarını) belirlenmesi amaçlandığı için betimsel ve deneysel bir araştırma türüdür. Deneysel araştırmada, iki ya da daha fazla grup üzerinde yapılan uygulamaların belli değişkenler açısından etkilerinin farklılaşma boyutu incelenir. Betimsel araştırma ise, bir konudaki mevcut durumu araştırmak ve belirlemektir.

3.2. Çalışma Grubu

Araştırmanın ana unsuru olan ROC analizinin yönetsel kısmının uygulanacağı klinik vakaların; var olan mevcut bilgiler ve yapılan kısa araştırma sonucunda çalışmaya dahil edilecek parametreleri belirlenmiştir. Adıyaman Üniversitesi Eğitim ve Araştırma Hastanesi İç Hastalıkları Polikliniğine belirtilen klinik vaka şüphesi ile giriş yapmış hastalar çalışma gruplarımız olup, kayıtlı olan bilgileri kullanılmıştır. Arşiv taraması esnasında, çalışmada kullanılacak parametrelerin sonuçlarının eksik olması dışlanma kriterlerimizden birisi olmuştur. Çalışma kapsamında belirtilen yöntemi uygulayacağımız *diyabet ve akut pankreatit klinik vakalarının* denek sayıları sırasıyla 311 ve 306'dır. Çalışmada yer alan hastaların cinsiyetlerine göre dağılımları Tablo 7 ve Tablo 8'de verilmiştir.

Tablo 7. Diyabet vakası cinsiyet dağılımları

Cinsiyet	N	%
Kadın	218	70,1
Erkek	93	29,9

Tablo 8. Akut pankreatit vakası cinsiyet dağılımları

Cinsiyet	N	%
Kadın	222	72,5
Erkek	84	27,5

3.3. Verilerin Toplanması

Araştırmanın verileri Adıyaman Üniversitesi Eğitim ve Araştırma Hastanesi İç hastalıkları Polikliniğine diyabet ve akut pankreatit şüphesi ile giriş yapmış hastaların dosya taramaları neticesinde kayıtlı bilgileri üzerinden retrospektif olarak alınmıştır.

Araştırma da kullanılacak veriler Adıyaman Üniversitesi Eğitim ve Araştırma Hastanesi Başhekimliğinin ve İç Hastalıkları Anabilim Dalı başkanlığının onayı ve de Süleyman Demirel Üniversitesi Tıp Fakültesi Etik Kurul onay kararı ile çalışmaya dahil edilmiştir.

3.4. Verilerin Analizi

3.4.1. Diyabet

Diabetes Mellitus (DM) insülin hormon sekresyonunun ve/veya insülin etkisinin mutlak veya göreceli eksikliği sonucu karbonhidrat, lipit ve protein metabolizmasında bozukluklara yol açan kronik hiperglisemi ile karakterize etiolojisinde birden fazla faktörün rol oynadığı metabolik bir hastalıktır (31). Tüm

dünyada morbitide ve mortalitenin en önemli nedenlerinden biri olan tip 2 diyabet artan obezite, fiziksel inaktivite ve nüfusun yaşlanması gibi nedenlerle gittikçe artmaktadır. Dünya sağlık örgütü (WHO) başta olmak üzere tüm nüfuslar arası kuruluşlar diyabet tanısı için doğrudan plazma glukoz ölçümüne dayanan tanı kriterleri olarak;

- Rastgele glukoz (+diyabet semptomları) ≥ 200 mg/dl
- APG (açlık plazma glukoz) (en az 8 saatlik açlığı takiben) ≥ 126 mg/dl
- OGTT'de 2. saat plazma glukoz ≥ 200 mg/dl
- Bozulmuş açlık glukozu : OGTT'de 2. saat plazma glukoz 160-199 mg/dl arası
- Bozulmuş glukoz toleransı (en az 8 saatlik açlığı takiben) 100-125 mg/dl arası

ortaya koymuştur (32).

Altın Standart

Hastalık durumu; 75 mg OGTT'nin 2. saatindeki plazma glukoz seviyesinin 200 mg/dl üstünde veya altında olması sonucuna göre "Hasta (Diyabet)" ve "Sağlıklı" şeklinde iki sınıflı olarak belirlenir. Diyabet kesin tanısı konulurken OGTT 2. saat değeri kullanılarak karar verilir. Eğer OGTT değeri yukarıdaki belirtilen kritere göre diyabet kategorisine giriyor ise bireye "Diyabet" tanısı konulması için yeterlidir. Burada "Sağlıklı" olarak tanımlanan grup diyabeti olmayan grup olarak olarak tanımlanmıştır.

Tanı Testleri

Bu çalışmada, Hemogloblin A1c (HbA1c), glukoz (açlık kan şekeri), üre ve kreatinin diyabet tanısında tanı testi olarak kabul edilmiş ve tanı performansları 75 mg OGTT 2. saat plazma glukoz seviyesinin sonucuna göre değerlendirilmiştir. HbA1c testi kandaki hemoglobline yapışan şeker (glukoz) değerlerini ölçer. Bu test genel olarak HbA1c %5,7 ise normal, %5,7 ile %6,4 ön diyabet ve %6,5'den fazla ise diyabet durumunu gösterir.

3.4.2. Akut Pankreatit

Akut pankreatit, pankreas enzimlerinin (amilaz, lipaz) çeşitli nedenlerle pankreas içinde aktifleşmesi ve pankreas dokusunun kendisini sindirmesi ile karakterize olan, serum pankreatik enzim düzeylerini yükselten, karın ağrısı ile karakterize özellik gösteren, mortalite ve morbiditesi yüksek inflamatuvar bir hastalıktır. Akut pankreatit tanı kriterleri olarak;

- Akut pankreatit için karakteristik tarzda karın ağrısı olması (ani başlangıçlı kuşak tarzında üst karın ağrısı)
- Serum amilaz ve/veya lipaz değerlerinin normalin üç kat veya daha fazla olması
- Ultrasonografi veya bilgisayarlı tomografi’de akut pankreatit için spesifik bulguların olması

ortaya konmuştur (33).

Altın Standart

Hastalık durumu, bilgisayarlı tomografi (BT) sonucuna göre “Sağlıklı” ve “Hasta (akut pankreatit)” şeklinde iki sınıflı olarak belirlenir. Akut pankreatit kesin tanısı konulurken BT sonuçları kullanılarak karar verilir.

Tanı Testleri

Bu çalışmada, lipaz ve amilaz enzimleri akut pankreatit tanısında tanı testi olarak kabul edilmiş ve tanı performansları BT sonucuna göre değerlendirilmiştir.

En iyi kesim noktalarının (cutoff) belirlenmesi, eğri altında alanın hesaplanması ve standart hatanın kestirimi için MedCalc 9.2.0.1 programı kullanılarak ROC analizi ile yapılmıştır. Parametrik olmayan ve yarı parametrik yaklaşımlara göre bootstrap (trapezoidal), Mann Whitney-U istatistik, Wilcoxon istatistik ve uygun binormal model yöntemleri ile AUC ve standart hata kestirimi için Metz ROC yazılımı çok platformlu Java GUI ROC 1.0.3 uygulaması yardımıyla yapılmıştır. LABROC, ROCKIT ve LABMRMC yazılımlarından LABROC kullanılmıştır.

4. BULGULAR

Bu bölümde, ele aldığımız ROC analizinin araştırmamızda yer alan klinik vakalar üzerinde uygulaması yapılmış ve elde edilen bulgular ile yorumlara yer verilmiştir.

4.1. Diyabet

Tablo 9 incelendiğinde çalışmaya 311 kişi dahil edilmiştir. Çalışmaya dahil edilen 311 kişiye diyabet şüphesiyle 75 mg şeker yüklemesi (OGTT) ile HbA1c, glukoz (açlık kan şekeri), üre, kreatinin biyokimya laboratuvar sonuçları istenmiştir. Diyabet şüphesi ile 75 mg OGTT yapılan 311 kişiden 2 saat sonunda çıkan sonuçlar neticesinde 200 ve üzeri değere sahip olanlar hasta grubunda (%26,37), 200 altı değere sahip olanlar ise sağlıklı grubunda (%73,63) yer almıştır.

Tablo 9. OGTT Sonucu Hasta ve Sağlıklı Bireylerin Dağılımları

	N	%
Hasta	82	26,37
Sağlıklı	229	73,63
Toplam Birey Sayısı	311	100

Bu bilgilerden yola çıkarak, diyabet klinik vakasında HbA1c, glukoz, üre ve kreatinin tanı testlerinin ROC analizi ile optimum eşik değerleri (cutoff) ve test performans değerlendirmeleri sırasıyla yapılmıştır.

Burada bir hatırlatma yapılacak olursa, çalışmamızda ROC analizi için ön diyabet (prediabetes) seviyesini almadan direkt olarak normal veya diyabet durumundaki eşik değerinin (cutoff) tahmini yapılmıştır.

Tablo 10'da HbA1c tanı testi için ROC analiz sonuçları verilmiştir. Her bir eşik değerine karşılık gelen duyarlılık (TPR), seçicilik (TNR), pozitif olabilirlik oranı (LR⁺) ve negatif olabilirlik oranı (LR⁻) ölçütleri yer almaktadır. Elde edilen sonuçlar neticesinde ikiliden elde edilen en yüksek (max) duyarlılık ve seçicilik değerlerine

sahip olan satırdaki deęerler, alıřma konusu olan o test iin eřik deęerini (cutoff) verecektir.

Tablo 10. HbA1c Testi iin ROC Analizi Sonuları

Eřik deęeri	TPR Duyarlılık	TNR Seicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
4	100,00	0,00	1,00	0,00
4	100,00	0,44	1,00	0,00
4,1	100,00	1,31	1,01	0,00
4,2	100,00	1,75	1,02	0,00
4,3	100,00	2,18	1,02	0,00
4,4	100,00	3,06	1,03	0,00
4,5	100,00	3,93	1,04	0,00
4,6	100,00	4,80	1,05	0,00
4,8	100,00	5,68	1,06	0,00
4,9	100,00	8,73	1,10	0,00
5	98,78	11,35	1,11	0,11
:	:	:	:	:
6	78,05	56,77	1,81	0,39
6,1	71,95	62,88	1,94	0,45
6,2	69,51	67,25	2,12	0,45
6,3	62,20	71,62	2,19	0,53
6,4	54,88	77,73	2,46	0,58
:	:	:	:	:
9,4	4,88	98,25	2,79	0,97
9,7	4,88	98,69	3,72	0,96
10,4	3,66	98,69	2,79	0,98
11,7	2,44	98,69	1,86	0,99
11,9	1,22	98,69	0,93	1,00
12,4	1,22	99,13	1,40	1,00
12,7	0,00	99,13	0,00	1,00
13,3	0,00	99,56	0,00	1,00
17,7	0,00	100,00	0,00	1,00

HbA1c testi iin %6,2 deęerine karřılık gelen duyarlılık deęeri 69,51 ve seicilik deęeri 67,25 olup duyarlılık ve seicilik deęerlerinin toplamlarının en yksek olduęu 6,2 deęeri en uygun sınıflandırmayı yapan eřik (cutoff) deęerdir.

HbA1c lm yapılan bireylerde diyabet olup olmadıęını belirlemek iin temel alacaęımız referans deęer 6,2 olarak belirlenmiřtir. Bulunan duyarlılık lt deęeri ise HbA1c tanı testinin %69,51 ihtimalle gerek durumda diyabeti olan bireyleri doęru tespit edebildięini, seicilik deęeri ise HbA1c tanı testinin %67,2

ihtimalle gerçek durumda diyabet olmayan bireyleri doğru tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığı zaman HbA1c testinin diyabet olanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 10’de yer alan olabilirlik oranlarını değerlendirecek olursak elde edilen $2,12 LR^+$, diyabet olan bir bireyde HbA1c testi sonucunun pozitif çıkma olasılığının diyabet olmayan bir bireyde HbA1c testi sonucunun pozitif çıkma olasılığından 2,12 kat daha fazla olduğunu ifade etmektedir. $0,45 LR^-$ ise, diyabet olan bir bireyde HbA1c testi sonucunun negatif çıkma olasılığının diyabet olmayan bir bireyde HbA1c testi sonucunun negatif çıkma olasılığından 0,45 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu HbA1c testi için belirlenen 6,2 kesme puanına göre ikili atama tablosu Tablo 11’de verilmiştir. Tablo 11 incelendiğinde TP ve FP değerlerinin sırasıyla 59 ve 85 olduğu, FN ve TN değerlerinin ise sırasıyla 23 ve 144 olduğu görülmektedir.

Tablo 11. HbA1c Testinde 6,2 Eşik Değerine Göre İkili Atama Tablosu

	OGTT Sonucu (Gerçek Durum)		
	Diyabet	Sağlıklı	Toplam
HbA1c Tanı Testi Sonucu			
Pozitif	59	85	144
Negatif	23	144	167
Toplam	82	229	311

Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,41 ve NPV ise 0,86 bulunmuştur. Bulunan bu değerler sırasıyla, HbA1c testinin sonucu pozitif ise 0,41 olasılıkla gerçek durumda diyabet olduğunu, negatif ise 0,86 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Çalışmamızın içeriğinde bahsettiğimiz üzere ROC analizi optimum eşik değeri (cutoff) belirleme fonksiyonuna sahip olduğu gibi elde edilen bazı göstergeler

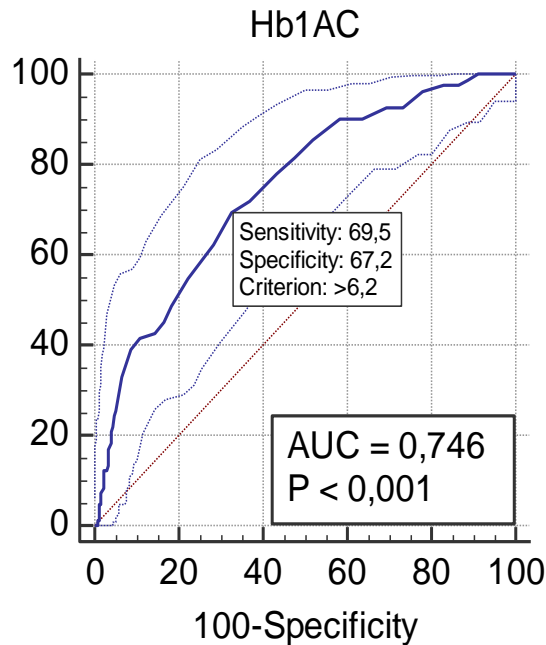
ile de tanı testlerinin performanslarını değerlendirebilmektedir. Eğri altında kalan alan hesabı ve TPR ile FPR arasındaki maksimum farkın hesabı tanı testimizin kestirim gücünü ve belirlenen eşik değerinin gerçek ile örtüşme yüzdesi hakkında bilgiler vermektedir.

HbA1c tanı testi için belirlenen eşik değerinde YI değerimiz 0,367 bulunmuştur. 0,367 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

HbA1c testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,746 ve GA (0,694-0,793) bulunmuş olup, kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir. İlâveten, p değeri de oldukça düşük bir değer olarak hesaplandığından HbA1c test ayırım sonucunun kabul edilebilir bir güç değerinde olduğunu teyit eder ($p < 0,001$).

Tablo 12. HbA1c Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	p
0,746	0,03	0,694	0,793	<0,001



Grafik 1. HbA1c Testi için ROC Eğrisi (%95 CI)

ROC eğrisi altında kalan alan parametrik yaklaşımlara göre hesaplanabildiği gibi çalışmamızda da belirttiğimiz gibi parametrik olmayan ve yarı parametrik olan yaklaşımlara göre de hesaplanabilmektedir. Bu yaklaşımlara göre olan hesaplamalarımızı Metz ROC yazılımı çok platformlu Java GUI ROC 1.0.3 uygulaması yardımıyla yapılmıştır. LABROC, ROCKIT ve LABMRMC yazılımlarından LABROC kullanılmıştır.

LABROC yazılımı yardımıyla HbA1c testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis), U-istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralıkları Tablo 13’de verilmiştir.

Tablo 13 incelendiğinde parametrik olmayan üç yonteme göre AUC kestirimlerinin, standart hatalarının ve de güven aralıklarının aynı olduğu görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak hesaplanan AUC değerleri de aynı olup 0,746 bulunmuştur.

Tablo 13. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım (Trapezoidal, Mann-Whitney U, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,746	0,03	0,685	0,806
U-istatistik (Zhou and Gatsonis)	0,746	0,03	0,685	0,806
U-istatistik (Gallas)	0,746	0,03	0,685	0,806

LABROC yazılımı yardımıyla HbA1c testinin sonuçları yarı parametrik yaklaşıma ve uygun binormal modele göre hesaplanan eğri altında kalan alan, standart hata Tablo 14’de verilmiştir.

MedCalc ve LABROC yazılımları kullanılarak hesaplanan AUC kestirimleri ve standart hata arasında farklılık olduğu görülmüş olup her iki değerde MedCalc’a göre çok küçük farklılıklar neticesinde daha küçük çıkmıştır.

Tablo 14. Yarı Parametrik Yaklaşım Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım (Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,744	0,029

Tablo 15’de glukoz tanı testi için ROC analiz sonuçları verilmiştir.

Tablo 15. Glukoz Testi için ROC Analizi Sonuçları

Eşik değeri	TPR Duyarlılık	TNR Seçicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
3	100,00	0,00	1,00	0,00
3	100,00	0,44	1,00	0,00
64	100,00	0,87	1,01	0,00
67	100,00	1,31	1,01	0,00
68	100,00	1,75	1,02	0,00
76	100,00	2,18	1,02	0,00
79	100,00	2,62	1,03	0,00
80	100,00	3,93	1,04	0,00
82	100,00	6,11	1,07	0,00
84	100,00	8,73	1,10	0,00
86	100,00	9,61	1,11	0,00
87	100,00	11,35	1,13	0,00
88	97,56	12,66	1,12	0,19
:	:	:	:	:
115	68,29	69,43	2,23	0,46
116	67,07	72,49	2,44	0,45
117	59,76	73,36	2,24	0,55
:	:	:	:	:
198	7,32	98,25	4,19	0,94
208	6,10	98,25	3,49	0,96
210	4,88	98,25	2,79	0,97
211	4,88	98,69	3,72	0,96
248	4,88	99,13	5,59	0,96
262	4,88	99,56	11,17	0,96
270	2,44	100,00	11,17	0,98
286	1,22	100,00	11,17	0,99
323	0,00	100,00	11,17	1,00

Glukoz testi için 116 değerine karşılık gelen duyarlılık değeri 67,07 ve seçicilik değeri 72,49 olup duyarlılık ve seçicilik değerlerinin toplamlarının en yüksek olduğu 116 değeri en uygun sınıflandırmayı yapan eşik (cutoff) değeridir.

Glukoz ölçümü yapılan bireylerde diyabet olup olmadığını belirlemek için temel alacağımız referans değer 116 olarak belirlenmiştir. Bulunan duyarlılık ölçüt değeri ise glukoz tanı testinin %67,07 ihtimalle gerçek durumda diyabeti olan bireyleri doğru tespit edebildiğini, seçicilik değeri ise glukoz tanı testinin %72,49 ihtimalle gerçek durumda diyabet olmayan bireyleri doğru tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığı zaman glukoz testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 15’de yer alan olabilirlik oranlarını değerlendirecek olursak elde edilen $2,44 LR^+$, diyabet olan bir bireyde glukoz testi sonucunun pozitif çıkma olasılığının diyabet olmayan bir bireyde glukoz testi sonucunun pozitif çıkma olasılığından 2,44 kat daha fazla olduğunu ifade etmektedir. $0,45 LR^-$ ise, diyabet olan bir bireyde glukoz testi sonucunun negatif çıkma olasılığının diyabet olmayan bir bireyde glukoz testi sonucunun negatif çıkma olasılığından 0,45 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu glukoz testi için belirlenen 116 kesme puanına göre ikili atama tablosu Tablo 16’de verilmiştir. Tablo 16 incelendiğinde TP ve FP değerlerinin sırasıyla 56 ve 70 olduğu, FN ve TN değerlerinin ise sırasıyla 26 ve 159 olduğu görülmektedir.

Tablo 16. Glukoz Testinde 116 Eşik Değerine Göre İkili Atama Tablosu

Glukoz Tanı Testi Sonucu	OGTT Sonucu (Gerçek Durum)		
	Diyabet	Sağlıklı	Toplam
	Pozitif	56	70
Negatif	26	159	185
Toplam	82	229	311

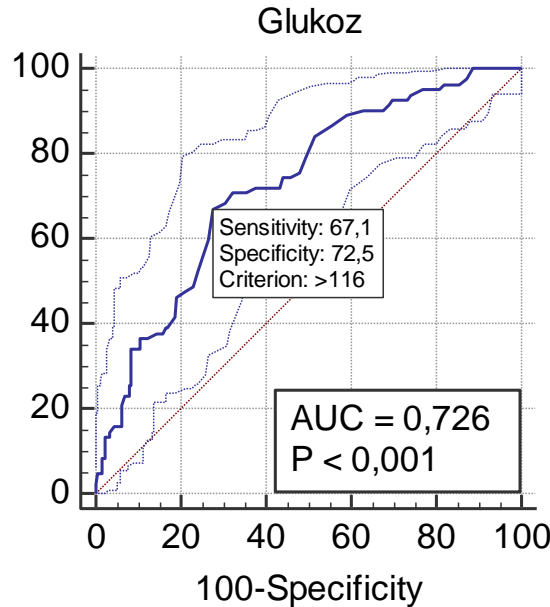
Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,44 ve NPV ise 0,85 bulunmuştur. Bulunan bu değerler sırasıyla, glukoz testinin sonucu pozitif ise 0,44 olasılıkla gerçek durumda diyabet olduğunu, negatif ise 0,85 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Glukoz tanı testi için belirlenen eşik değerinde YI değerimiz 0,395 bulunmuştur. 0,395 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

Glukoz testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,726 ve GA (0,673-0,775) bulunmuş olup, kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir. İlâveten, p değeri de oldukça düşük bir değer olarak hesaplandığından glukoz test ayırım sonucunun kabul edilebilir bir güç değerinde olduğunu teyit eder ($p < 0,001$).

Tablo 17. Glukoz Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	p
0,726	0,031	0,673	0,775	<0,001



Grafik 2. Glukoz Testi için ROC Eğrisi (%95 CI)

LABROC yazılımı yardımıyla glukoz testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis), U-istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralıkları Tablo 18’de verilmiştir.

Tablo 18 incelendiğinde parametrik olmayan üç yönteme göre AUC kestirimlerinin, standart hatalarının aynı olduğu, güven aralıklarının ise küçük farklılıklar gösterdiği görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak hesaplanan AUC değerleri de aynı olup 0,726 bulunmuştur.

Yarı parametrik yaklaşıma göre ise AUC kestirimleri arasında farklılık olduğu görülmüş olup MedCalc’a göre çok küçük farklılıklar neticesinde daha büyük çıkmıştır.

Tablo 18. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım (Trapezoidal, Mann-Whitney U, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,726	0,029	0,667	0,784
U-istatistik (Zhou and Gatsonis)	0,726	0,031	0,663	0,788
U-istatistik (Gallas)	0,726	0,031	0,664	0,787

Tablo 19. Yarı Parametrik Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım (Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,727	0,031

Tablo 20’de üre tanı testi için ROC analiz sonuçları verilmiştir.

Tablo 20. Üre Testi için ROC Analizi Sonuçları

Eşik değeri	TPR Duyarlılık	TNR Seçicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
10	100,00	0,00	1,00	0,00
10	98,78	0,00	0,99	0,00
11	98,78	0,44	0,99	2,79
12	97,56	0,87	0,98	2,79
13	97,56	1,75	0,99	1,4
14	97,56	3,06	1,01	0,8
15	96,34	5,24	1,02	0,7
16	92,68	6,99	1,00	1,05
17	91,46	10,48	1,02	0,81
18	89,02	13,97	1,03	0,79
19	82,93	17,03	1,00	1,00
:	:	:	:	:
26	57,32	49,78	1,14	0,86
27	51,22	54,15	1,12	0,9
28	46,34	59,83	1,15	0,9
29	41,46	65,94	1,22	0,89
30	35,37	69,43	1,16	0,93
31	30,49	72,49	1,11	0,96
32	26,83	75,11	1,08	0,97
33	23,17	77,29	1,02	0,99
:	:	:	:	:
85	0,00	98,69	0,00	1,01
126	0,00	99,13	0,00	1,01
127	0,00	99,56	0,00	0,00
132	0,00	100,00	0,00	0,00

Üre testi için 29 değerine karşılık gelen duyarlılık değeri 41,46 ve seçicilik değeri 65,94 olup duyarlılık ve seçicilik değerlerinin toplamının en yüksek olduğu 29 değeri en uygun sınıflandırmayı yapan eşik (cutoff) değeridir. Genel olarak üre testleri erkeklerde $31,8 \pm 65$ mg/100ml ve kadınlarda ise $28,7 \pm 7,1$ mg/100 ml değerleri referans olarak kabul edilmektedir.

Üre ölçümü yapılan bireylerde diyabet olup olmadığını belirlemek için temel alacağımız referans değer 29 olarak belirlenmiştir. Bulunan duyarlılık ölçüt değeri ise üre tanı testinin %41,46 ihtimalle gerçek durumda diyabeti olan bireyleri doğru tespit edebildiğini, seçicilik değeri ise üre tanı testinin %65,94 ihtimalle gerçek

durumda diyabet olmayan bireyleri doğru tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığında zaman üre testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 20’de yer alan olabilirlik oranlarını değerlendirecek olursak elde edilen 1,22 LR⁺, diyabet olan bir bireyde üre testi sonucunun pozitif çıkma olasılığının diyabet olmayan bir bireyde üre testi sonucunun pozitif çıkma olasılığından 1,22 kat daha fazla olduğunu ifade etmektedir. 0,89 LR⁻ ise; diyabet olan bir bireyde üre testi sonucunun negatif çıkma olasılığının diyabet olmayan bir bireyde üre testi sonucunun negatif çıkma olasılığından 0,89 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu üre testi için belirlenen 29 kesme puanına göre ikili atama tablosu Tablo 21’de verilmiştir. Tablo 21 incelendiğinde TP ve FP değerlerinin sırasıyla 38 ve 92 olduğu, FN ve TN değerlerinin ise sırasıyla 44 ve 137 olduğu görülmektedir.

Tablo 21. Üre Testinde 29 Eşik Değerine Göre İkili Atama Tablosu

	OGTT Sonucu (Gerçek Durum)		
	Diyabet	Sağlıklı	Toplam
	Üre Tanı Testi Sonucu		
Pozitif	38	92	130
Negatif	44	137	181
Toplam	82	229	311

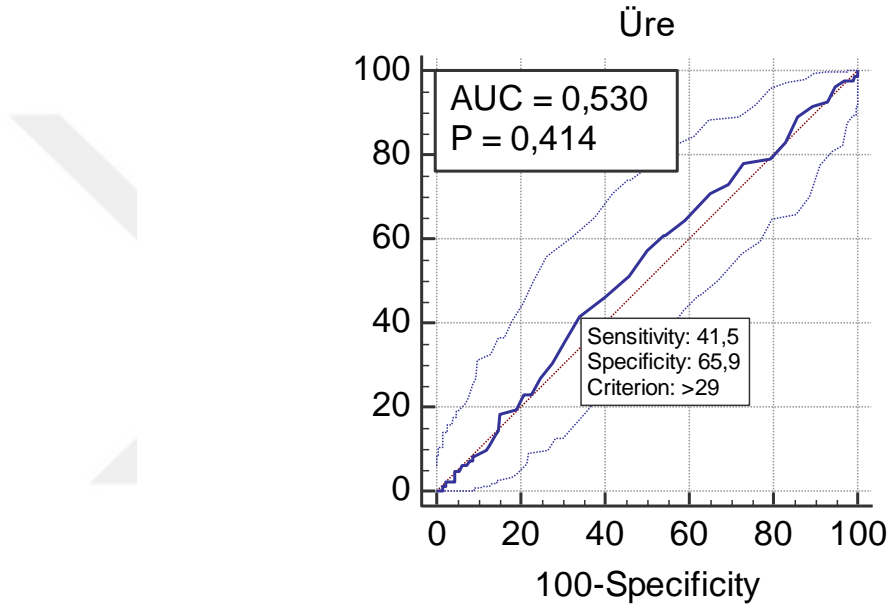
Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,29 ve NPV ise 0,75 bulunmuştur. Bulunan bu değerler sırasıyla, üre testinin sonucu pozitif ise 0,29 olasılıkla gerçek durumda diyabet olduğunu, negatif ise 0,75 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Üre tanı testi için belirlenen eşik değerinde YI değerimiz 0,074 bulunmuştur. 0,074 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

Üre testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,53 ve GA (0,473-0,587) bulunmuş olup, zayıf bir ayırım gücüne sahip olduğu gözlenmiştir. İlaveeten, bulunan p değeri de bir anlamlılık ifade etmemektedir ($p>0,05$).

Tablo 22. Üre Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	P
0,530	0,037	0,473	0,587	0,414



Grafik 3. Üre Testi için ROC Eğrisi (%95 CI)

LABROC yazılımı yardımıyla üre testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis), U-istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralıkları Tablo 23’de verilmiştir.

Tablo 23 incelendiğinde parametrik olmayan üç yönteme göre AUC kestirimlerinin, standart hatalarının aynı olduğu görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak parametrik olmayan yaklaşıma göre hesaplanan AUC değerleri de aynı olup 0,53 bulunmuştur.

Yarı parametrik yaklaşıma göre ise, AUC kestirimi küçük bir farkla daha küçük çıkmış olup 0,519 bulunmuştur.

Tablo 23. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım (Trapezoidal, Mann-Whitney U, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,530	0,037	0,459	0,604
U-istatistik (Zhou and Gatsonis)	0,530	0,037	0,457	0,602
U-istatistik (Gallas)	0,530	0,037	0,457	0,602

Tablo 24. Yarı Parametrik Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım (Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,519	0,036

Tablo 25’de kreatinin tanı testi için ROC analiz sonuçları verilmiştir.

Tablo 25. Kreatinin Testi için ROC Analizi Sonuçları

Eşik değeri	TPR Duyarlılık	TNR Seçicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
0,5	0,00	100,00	0,00	1,00
0,5	2,44	96,94	0,80	1,01
0,51	3,66	96,51	1,05	1,00
0,53	4,88	95,63	1,12	0,99
0,55	6,10	95,63	1,40	0,98
0,56	6,10	95,20	1,27	0,99
0,57	6,10	94,76	1,16	0,99
0,58	7,32	93,89	1,20	0,99
0,59	7,32	93,45	1,12	0,99
:	:	:	:	:
0,76	56,10	44,54	1,01	0,99
0,77	57,32	44,10	1,03	0,97
0,78	59,76	43,23	1,05	0,93
0,79	63,41	42,79	1,11	0,85
0,8	74,39	31,88	1,09	0,80
0,81	74,39	30,57	1,07	0,84
0,82	76,83	27,95	1,07	0,83
0,83	76,83	26,64	1,05	0,87
0,84	79,27	25,76	1,07	0,80
:	:	:	:	:
1,73	97,56	1,31	0,99	1,86
1,79	97,56	0,87	0,98	2,79
2	98,78	0,44	0,99	2,79
6,4	98,78	0,00	0,99	0,00
8	100,00	0,00	1,00	0,00

Kreatinin testi için 0,8 değerine karşılık gelen duyarlılık değeri 74,39 ve seçicilik değeri 31,88 olup duyarlılık ve seçicilik değerlerinin toplamalarının en yüksek olduğu 0,8 değeri en uygun sınıflandırmayı yapan eşik (cutoff) değeridir. Genel olarak kreatin değerleri erkeklerde 0,5 ile 1,2 ve kadınlarda ise 0,5 ile 1,1 mg/dL arasında normal değerler olarak kabul edilmektedir.

Kreatinin ölçümü yapılan bireylerde diyabet olup olmadığını belirlemek için temel alacağımız referans değer 0,8 olarak belirlenmiştir. Bulunan duyarlılık ölçüt değeri ise kreatinin tanı testinin %74,39 ihtimalle gerçek durumda diyabeti olan bireyleri doğru tespit edebildiğini, seçicilik değeri ise kreatinin tanı testinin %31,88 ihtimalle gerçek durumda diyabet olmayan bireyleri doğru tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığı zaman kreatinin testinin diyabet olanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 25’da yer alan olabilirlik oranlarını değerlendirecek olursak, elde edilen 1,09 LR⁺, diyabet olan bir bireyde kreatinin testi sonucunun pozitif çıkma olasılığının, diyabet olmayan bir bireyde kreatinin testi sonucunun pozitif çıkma olasılığından 1,09 kat daha fazla olduğunu ifade etmektedir. 0,8 LR⁻ ise, diyabet olan bir bireyde kreatinin testi sonucunun negatif çıkma olasılığının, diyabet olmayan bir bireyde kreatinin testi sonucunun negatif çıkma olasılığından 0,8 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu kreatinin testi için belirlenen 0,8 kesme puanına göre ikili atama tablosu Tablo 26’de verilmiştir. Tablo 26 incelendiğinde TP ve FP değerlerinin sırasıyla 30 ve 98 olduğu, FN ve TN değerlerinin ise sırasıyla 52 ve 131 olduğu görülmektedir.

Tablo 26. Kreatinin Testinde 0,8 Eşik Değerine Göre İkili Atama Tablosu

	OGTT Sonucu (Gerçek Durum)			
	Diyabet	Sağlıklı	Toplam	
Kreatinin Tanı Testi Sonucu	Pozitif	30	98	128
	Negatif	52	131	183
	Toplam	82	229	311

Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,23 ve NPV ise 0,71 bulunmuştur. Bulunan bu değerler sırasıyla, kreatinin testinin sonucu pozitif ise

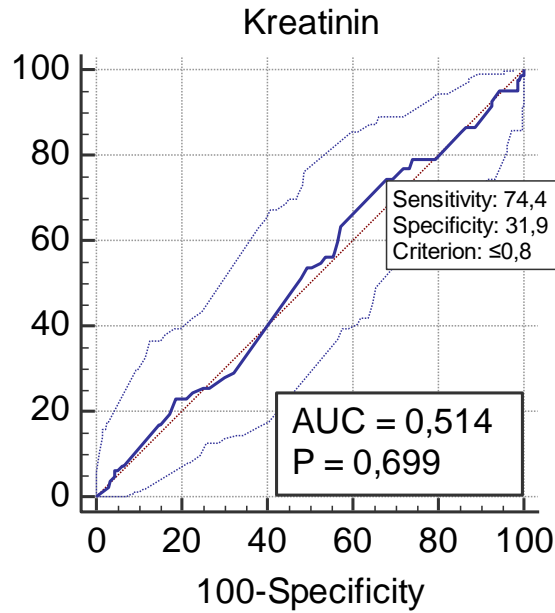
0,23 olasılıkla gerçek durumda diyabet olduğunu, negatif ise 0,71 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Kreatinin tanı testi için belirlenen eşik değerinde YI değerimiz 0,062 bulunmuştur. 0,062 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

Kreatinin testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,514 ve GA (0,457-0,571) bulunmuş olup, zayıf bir ayırım gücüne sahip olduğu gözlenmiştir. İlâveten, bulunan p değeri de bir anlamlılık ifade etmemektedir ($p>0,05$).

Tablo 27. Kreatinin Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	P
0,514	0,037	0,457	0,571	0,699



Grafik 4. Kreatinin Testi için ROC Eğrisi (%95 CI)

LABROC yazılımı yardımıyla kreatinin testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis),

U-istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralıkları Tablo 28’de verilmiştir.

Tablo 28 incelendiğinde parametrik olmayan üç yönteme göre AUC kestirimlerinin, standart hatalarının aynı olduğu, en geniş güven aralığının ise Bootstrap modelinde olduğu görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak parametrik olmayan yaklaşıma göre hesaplanan AUC değerleri aynı olmayıp, daha düşük bulunmuştur.

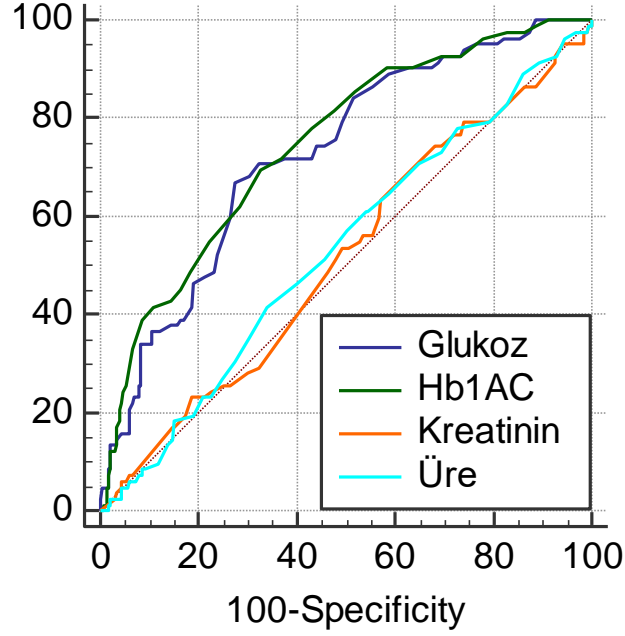
Yarı parametrik yaklaşıma göre kestirilen AUC değeri ise parametrik olmayan yaklaşıma göre daha yakın bir kestirim elde etmiştir.

Tablo 28. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım (Trapezoidal, Mann-Whitney U-statistic, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,485	0,037	0,411	0,559
U-istatistik (Zhou and Gatsonis)	0,485	0,037	0,412	0,558
U-istatistik (Gallas)	0,485	0,037	0,412	0,558

Tablo 29. Yarı Parametrik Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım (Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,508	0,029



Grafik 5. Diyabet için Tanı Testlerinin Performans Karşılaştırmaları

Diyabet klinik vakasında kullanılan testlerin performans değerlendirmeleri sonucu en iyi kestirimin HbA1c (AUC=0,746; $p<0,001$) ve glukoz (AUC=0,726; $p<0,001$) olduğu görülmüştür. Diğer yönden ise Kreatinin (AUC=0,530; $p>0,05$) ve üre (AUC=0,514; $p>0,05$) testlerinde ayırım gücü bakımından bir anlamlılık görülmemiştir.

4.2. Akut Pankreatit

Tablo 30 incelendiğinde çalışmaya 306 kişi dahil edilmiştir. Çalışmaya dahil edilen 306 kişiye akut pankreatit şüphesiyle alt ve üst abdomen bölgelere bilgisayarlı tomografi (BT) çekimi ile lipaz ve amilaz biyokimya laboratuvar sonuçları istenmiştir. Akut pankreatit şüphesi ile BT çekilen 306 kişiden elde edilen sonuçlara göre hasta grubunda 132 (%43,14) kişi, sağlıklı grubunda 174 (%56,86) kişi yer almıştır.

Tablo 30. BT Sonucu Hasta ve Sağlıklı Bireylerin Dağılımları

	N	%
Hasta	132	43,14
Sağlıklı	174	56,86
Toplam Birey Sayısı	306	100

Bu bilgilerden yola çıkarak, akut pankreatit klinik vakasında lipaz ve amilaz tanı testlerinin ROC analizi ile optimum eşik değerleri (cutoff) ve test performans değerlendirmeleri sırasıyla yapılmıştır.

Tablo 31. Lipaz Testi için ROC Analizi Sonuçları

Eşik değeri	TPR Duyarlılık	TNR Seçicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
6	100,00	0,00	1,00	0,00
6	97,73	0,00	0,98	0,00
7	97,73	6,90	1,05	0,33
8	93,18	6,90	1,00	0,99
11	93,18	10,34	1,04	0,66
12	93,18	13,79	1,08	0,49
13	93,18	15,52	1,10	0,44
16	93,18	17,24	1,13	0,4
17	93,18	18,97	1,15	0,36
:	:	:	:	:
75	65,91	77,59	2,94	0,44
76	65,91	81,03	3,48	0,42
78	65,91	82,76	3,82	0,41
82	65,91	84,48	4,25	0,4
83	63,64	86,21	4,61	0,42
84	61,36	86,21	4,45	0,45
88	59,09	87,93	4,90	0,47
90	59,09	89,66	5,71	0,46
:	:	:	:	:
629	15,91	98,28	9,23	0,86
902	13,64	98,28	7,91	0,88
1095	11,36	98,28	6,59	0,9
1200	0,00	100,00	0,00	1,00

Lipaz testi için 82 değerine karşılık gelen duyarlılık değeri 65,91 ve seçicilik değeri 84,48 olup duyarlılık ve seçicilik değerlerinin toplamalarının en yüksek olduğu 82 değeri en uygun sınıflandırmayı yapan eşik (cutoff) değeridir.

Lipaz ölçümü yapılan bireylerde akut pankreatit olup olmadığını belirlemek için temel alacağımız referans değer 82 olarak belirlenmiştir. Bulunan duyarlılık ölçüt değeri ise lipaz tanı testinin %65,91 ihtimalle gerçek durumda akut pankreatit olan bireyleri doğru tespit edebildiğini, seçicilik değeri ise lipaz tanı testinin %84,48 ihtimalle gerçek durumda akut pankreatit olmayan bireyleri doğru tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığı zaman lipaz testinin akut pankreatit olmayanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 31’de yer alan olabilirlik oranlarını değerlendirecek olursak, elde edilen $4,25 LR^+$, akut pankreatit olan bir bireyde lipaz testi sonucunun pozitif çıkma olasılığının akut pankreatit olmayan bir bireyde lipaz testi sonucunun pozitif çıkma olasılığından 4,25 kat daha fazla olduğunu ifade etmektedir. $0,4 LR^-$ ise, akut pankreatit olan bir bireyde lipaz testi sonucunun negatif çıkma olasılığının, akut pankreatit olmayan bir bireyde lipaz testi sonucunun negatif çıkma olasılığından 0,4 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu lipaz testi için belirlenen 82 kesme puanına göre ikili atama tablosu Tablo 32’de verilmiştir. Tablo 32 incelendiğinde TP ve FP değerlerinin sırasıyla 87 ve 30 olduğu, FN ve TN değerlerinin ise sırasıyla 45 ve 144 olduğu görülmektedir.

Tablo 32. Lipaz Testinde 82 Eşik Değerine Göre İkili Atama Tablosu

	BT Sonucu (Gerçek Durum)			
	Diyabet	Sağlıklı	Toplam	
Lipaz Tanı Testi Sonucu	Pozitif	87	30	117
	Negatif	45	144	189
	Toplam	132	174	311

Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,74 ve NPV ise 0,76 bulunmuştur. Bulunan bu değerler sırasıyla, lipaz testinin sonucu pozitif ise 0,74

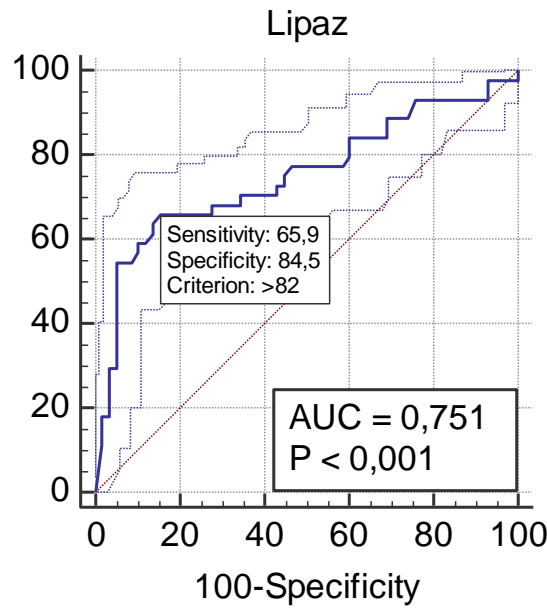
olasılıkla gerçek durumda akut pankreatit olduğunu, negatif ise 0,76 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Lipaz tanı testi için belirlenen eşik değerinde YI değerimiz 0,503 bulunmuştur. 0,503 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

Lipaz testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,751 ve GA (0,699-0,999) bulunmuş olup, kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir. İlâveten, p değeri de oldukça düşük bir değer olarak hesaplandığından lipaz test ayırım sonucunun kabul edilebilir bir güç değerinde olduğunu teyit eder ($p < 0,001$).

Tablo 33. Lipaz Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	P
0,751	0,03	0,699	0,999	<0,001



Grafik 6. Lipaz Testi için ROC Eğrisi

LABROC yazılımı yardımıyla lipaz testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis), U-

istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralıkları Tablo 34’de verilmiştir.

Tablo 34 incelendiğinde parametrik olmayan üç yönteme göre AUC kestirimlerinin aynı olduğu, en düşük standart hata ile en dar güven aralığına ise Bootstrap modelinin sahip olduğu görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak parametrik olmayan yaklaşıma göre hesaplanan AUC değerleri de aynı olup 0,751 bulunmuştur.

Yarı parametrik yaklaşıma göre elde edilen AUC kestirim değerinin aynı olduğu, standart hatanın ise küçük bir farkla daha düşük çıktığı görülmüştür.

Tablo 34. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım				
(Trapezoidal, Mann-Whitney U, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,751	0,029	0,693	0,808
U-istatistik (Zhou and Gatsonis)	0,751	0,03	0,692	0,810
U-istatistik (Gallas)	0,751	0,03	0,692	0,809

Tablo 35. Yarı Parametrik Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım			
(Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,751	0,028

Tablo 36’da amilaz tanı testi için ROC analiz sonuçları verilmiştir.

Tablo 36. Amilaz Testi için ROC Analizi Sonuçları

Eşik değeri	TPR Duyarlılık	TNR Seçicilik	Pozitif Olabilirlik Oranı	Negatif Olabilirlik Oranı
7	100,00	0,00	1,00	0,00
7	100,00	1,72	1,02	0,00
18	100,00	3,45	1,04	0,00
21	100,00	6,90	1,07	0,00
25	97,73	8,62	1,07	0,26
29	97,73	10,34	1,09	0,22
32	95,45	10,34	1,06	0,44
:	:	:	:	:
83	65,91	65,52	1,91	0,52
86	65,91	67,24	2,01	0,51
88	63,64	68,97	2,05	0,53
89	63,64	70,69	2,17	0,51
90	63,64	72,41	2,31	0,5
91	61,36	72,41	2,22	0,53
92	56,82	72,41	2,06	0,6
95	54,55	72,41	1,98	0,63
97	50,00	72,41	1,81	0,69
98	47,73	72,41	1,73	0,72
:	:	:	:	:
497	9,09	98,28	5,27	0,93
570	6,82	98,28	3,95	0,95
857	6,82	100,00	0,00	0,93
1080	4,55	100,00	0,00	0,95
1136	2,27	100,00	0,00	0,98
2634	0,00	100,00	0,00	1,00

Amilaz testi için 90 değerine karşılık gelen duyarlılık değeri 63,64 ve seçicilik değeri 72,41 olup duyarlılık ve seçicilik değerlerinin toplamının en yüksek olduğu 90 değeri en uygun sınıflandırmayı yapan eşik (cutoff) değeridir.

Amilaz ölçümü yapılan bireylerde akut pankreatit olup olmadığını belirlemek için temel alacağımız referans değer 90 olarak belirlenmiştir. Bulunan duyarlılık ölçüt değeri ise amilaz tanı testinin %63,64 ihtimalle gerçek durumda akut pankreatiti olan bireyleri doğru tespit edebildiğini, seçicilik değeri ise amilaz tanı testinin %72,41 ihtimalle gerçek durumda akut pankreatiti olmayan bireyleri doğru

tespit edebildiğini göstermektedir. Bulunan ihtimallere bakıldığı zaman amilaz testinin akut pankreatit olanları bulmakta daha başarılı olduğu ifade edilebilir.

Tablo 36’da yer alan olabilirlik oranlarını değerlendirecek olursak, elde edilen 2,31 LR⁺, akut pankreatit olan bir bireyde amilaz testi sonucunun pozitif çıkma olasılığının akut pankreatit olmayan bir bireyde amilaz testi sonucunun pozitif çıkma olasılığından 2,31 kat daha fazla olduğunu ifade etmektedir. 0,5 LR⁻ ise, akut pankreatit olan bir bireyde amilaz testi sonucunun negatif çıkma olasılığının, akut pankreatit olmayan bir bireyde amilaz testi sonucunun negatif çıkma olasılığından 0,5 kat daha fazla olduğunu ifade etmektedir.

ROC analizi sonucu amilaz testi için belirlenen 90 kesme puanına göre ikili atama tablosu Tablo 37’de verilmiştir. Tablo 37 incelendiğinde TP ve FP değerlerinin sırasıyla 84 ve 51 olduğu, FN ve TN değerlerinin ise sırasıyla 48 ve 123 olduğu görülmektedir.

Tablo 37. Amilaz Testinde 90 Eşik Değerine Göre İkili Atama Tablosu

	BT Sonucu (Gerçek Durum)		
	Diyabet	Sağlıklı	Toplam
Amilaz Tanı Testi Sonucu	Pozitif	84	135
	Negatif	48	171
	Toplam	132	311

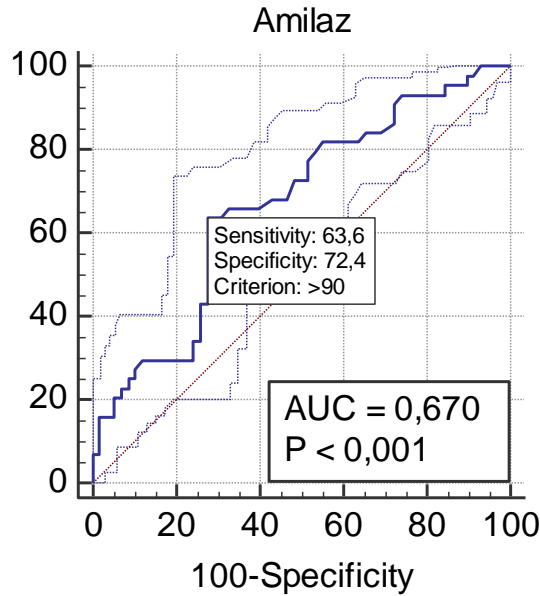
Yukarıda verilen TP, FP, FN ve TN değerlerinden yola çıkarak pozitif ve negatif tahmini değer ölçütlerinin hesaplanması neticesinde de PPV 0,62 ve NPV ise 0,71 bulunmuştur. Bulunan bu değerler sırasıyla, amilaz testinin sonucu pozitif ise 0,62 olasılıkla gerçek durumda akut pankreatit olduğunu, negatif ise 0,71 olasılıkla gerçek durumda sağlıklı olduğunu ifade etmektedir.

Amilaz tanı testi için belirlenen eşik değerinde YI değerimiz 0,36 bulunmuştur. 0,36 değerine karşılık gelen TPR ile FPR arasındaki fark en yüksektir.

Amilaz testinden elde edilen AUC değeri güven aralığı değeri ile birlikte 0,67 ve GA (0,614-0,722) bulunmuş olup, zayıf bir ayırım gücüne sahip olduğu gözlenmiştir. İlâveten, p değeri bir anlamlılık ifade etse de AUC değerinin 0,67 olması zayıf bir ayırım gücüne sahip olduğunu göstermektedir ($p < 0,001$).

Tablo 38. Amilaz Testi için ROC Eğrisi Altında Kalan Alan

AUC	S. Hata	% 95 Güven Aralığı Alt Sınır	% 95 Güven Aralığı Üst Sınır	P
0,670	0,030	0,614	0,722	<0,001



Grafik 7. Amilaz Testi için ROC Eğrisi

LABROC yazılımı yardımıyla amilaz testinin sonuçları parametrik olmayan yaklaşıma ve Bootstrap (Gruszauskas vd.), U-istatistik (Zhou and Gatsonis), U-istatistik (Gallas) modellerine göre hesaplanan eğri altında kalan alan, standart hata ve güven aralığı Tablo 39’de verilmiştir.

Tablo 39 incelendiğinde parametrik olmayan üç yönteme göre AUC kestirimlerinin aynı olduğu, standart hatanın ve güven aralıklarının çok küçük farkla yüksek olduğu görülmektedir.

MedCalc ve LABROC yazılımları kullanılarak parametrik olmayan yaklaşıma göre hesaplanan AUC değerleri birbirinden farklı çıkmıştır.

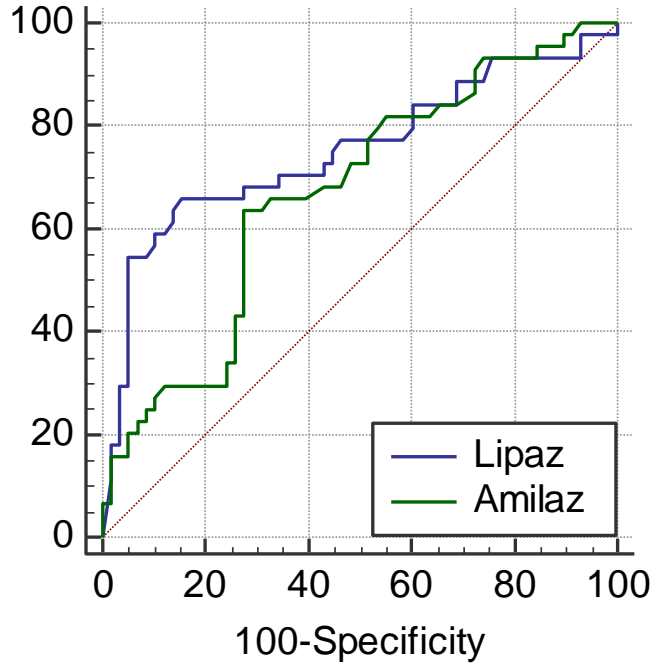
Yarı parametrik yaklaşıma göre ise; AUC kestirimi aynı olmayıp Medcalc'a göre küçük bir farkla daha yüksek olduğu görülmektedir.

Tablo 39. Parametrik Olmayan Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Parametrik Olmayan Yaklaşım (Trapezoidal, Mann-Whitney U, Wilcoxon Signed-Rank)				
Tahmin Yöntemi	AUC	S. Hata	% 95 Güven Sınırları	
			Alt Sınır	Üst Sınır
Bootstrap (Gruszauskas vd.)	0,669	0,032	0,607	0,732
U-istatistik (Zhou and Gatsonis)	0,669	0,031	0,609	0,730
U-istatistik (Gallas)	0,669	0,031	0,609	0,730

Tablo 40. Yarı Parametrik Yaklaşıma Göre ROC Eğrisi Altında Kalan Alan

Yarı Parametrik Yaklaşım (Conventional, Proper Binormal, Contaminated)			
Model	Tahmin Yöntemi	AUC	S. Hata
Uygun Binormal Model (PBM)	Bilgi matrisinin tersi (Inverse of Information Matrix)	0,677	0,030



Grafik 8. Akut Pankreatit için Tanı Testlerinin Performans Karşılaştırmaları

Akut pankreatit klinik vakasında kullanılan testlerin performans değerlendirmeleri sonucu en iyi kestirimin lipaz testi ($AUC=0,751$; $p<0,001$) olduğu görülmüştür. Diğer yönden ise amilaz ($AUC=0,670$; $p<0,001$) testi ayırım gücü bakımından bir anlamlılık ifade etse de zayıf bir ayırım gücüne sahip olduğu görülmektedir.

5. TARTIŞMA

Araştırma konusu olarak baz aldığımız ROC analizinin teorisi ve tekniği toplumda son zamanlarda sık görülen diyabet ve akut pankreatit klinik vakaları üzerinde uygulamalı bir şekilde ele alınmıştır. Günümüz toplumunda çok fazla kişide görülen ve ilerleyen yıllarda da artış göstereceği tahmin edilen diyabet tanısı tüm dünyada yüksek oranlarda morbidite ve mortaliteye sahip bir hastalıktır. Bu ve buna benzer hastalıklarda erken tanının önemi büyük olup, tanı koymada maliyetsiz, hızlı ve güvenilir yöntemlerin geliştirilmesi büyük önem taşımaktadır. Çalışmamızda ele aldığımız vakalar bakımından altın standart olarak uygulanan OGTT ve BT testleri maliyetli, uzun zaman alan, iş gücü gerektiren ve tekrarlanabilirliği düşük olan testlerdir. Diyabet tanısı için önemli olduğu kabul edilen biyokimyasal parametreler olan HbA1c, glukoz ve üre ile kreatinin, akut pankreatit tanısı içinde önemli olduğu kabul edilen biyokimyasal parametreler olan lipaz ve amilaz parametrelerinin tanı koymadaki performans değerlendirilmeleri ve her bir parametre için eşik değerleri (cutoff) ROC analizi ile incelenmiş ve belirlenmiştir.

Akarsu, 2008 yılında diyabet teşhisinde kullanılan testlerin tanısallık değerlendirilmeleri üzerine yaptığı çalışmasında (n=521) sağlıklı ve diyabetli grupların ayırımı için baz alınan parametrelerin ROC eğrilerinin karşılaştırılması yapılmış olup anlamlı bir farklılık elde edilmiştir ($p<0,05$). AUC değerleri ise HbA1c=0,853, glukoz=0,870 ve ürik asit=0,633 olarak bulunmuştur. HbA1c ve glukoz testlerinin AUC değerlerinin kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir ($p<0,001$). HbA1c testi için ROC analizi sonucuna göre en uygun eşik değer %6,1 olup duyarlılık %67,1, seçicilik %84,6, LR^+ 4,35, LR^- 0,39, PPV %76,3 ve NPV %77,7 olarak, glukoz testi için en uygun eşik değer 106 mg/dl olup duyarlılık %82,3, seçicilik %78,2, LR^+ 3,77, LR^- 0,23, PPV %71,3 ve NPV %87,0 olarak saptanmıştır. Akarsu çalışmasında farklı altın standartlar kullanmış olup altın standart olarak OGTT 120. dk. plazma glukozunu ele aldığı kısımda sağlıklı birey tayinini 140 mg/dl'in altı olarak yapmıştır (34).

Efe ve ark., 2014 yılında diyabet üzerine yaptıkları çalışmalarında (n=244) OGTT 120. dk. plazma glukozu baz alınarak bireyler sağlıklı ve diyabet olarak

gruplandırılmıştır. OGTT 2. saat plazma glukozuna göre HbA1c testi için ROC analizi sonucuna göre en iyi kesim noktası %6,35 olup duyarlılık %65,4, seçicilik %95, PPV %60,7, NPV %95,8 olarak bulunmuştur. Sağlıklı ve diyabetli grupların ayrımı için ele alınan HbA1c testinin ROC eğrisinden elde edilen AUC 0,769 değerinin (0,635-0,902) GA ile kabul edilebilir bir ayırım gücüne sahip olduğu bulunmuştur ($p<0,001$) (35).

Colagiuri ve ark., 2004 yılında Avusturalya bölgesinde yapılan çalışmalarında ($n=10447$) HbA1c testi için ROC analizi sonucuna göre en iyi kesim noktası %5,3 olup duyarlılık %78,7, seçicilik %82,8, LR^+ 4,58, LR^- 0,26 olarak saptanmıştır (36).

Mannucci ve ark., 2003 yılında İtalya bölgesinde yapılan çalışmalarında ($n=1215$) HbA1c testi için ROC analizi sonuçlarına göre en iyi eşik değeri %6.6 olup erkeklerde duyarlılık %98, kadınlarda ise %100, erkeklerde seçicilik %30, kadınlarda %21, erkeklerde LR^+ %1,3, kadınlarda LR^+ 1,28, erkeklerde LR^- 0,31, kadınlarda 0 olarak bulunmuştur (37).

Tavintharan ve ark., 2000 yılında diyabet üzerine yaptıkları çalışmalarında ($n=111$) HbA1c testi için en iyi kesim değeri %5,9 olup sensitivite %95, spesifite %67, LR^+ 2,88 ve LR^- 0,07 olarak bulunmuştur (38).

Çalışmamızda HbA1c testi için ROC analizi sonuçlarına göre en iyi kesim noktasını %6,2 olup duyarlılık %69,51, seçicilik %67,25, LR^+ 2,12, LR^- 0,45, PPV 0,41 ve NPV 0,86 olarak bulunmuştur. HbA1c testinden elde edilen AUC 0,751 değerinin (0,694-0,793) GA ile bulunup bunun kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir ($p<0,001$). Glukoz testi için ROC analizi sonuçlarına göre en iyi kesim noktası 116 mg/dl olup duyarlılık %67,07, seçicilik %72,49, LR^+ 2,44, LR^- 0,45, PPV 0,44 ve NPV 0,85 olarak bulunmuştur. Glukoz testinden elde edilen AUC 0,726 değerinin (0,673-0,775) GA ile bulunup bunun kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir ($p<0,001$). Yapılan çalışmaların var olan testlerine ait ROC analizi ölçütlerine bakıldığı zaman benzerlikler taşıdığı görülmektedir. Ancak en uygun eşik değerlerin belirlenmesinde ve kullanılan ölçütlerin saptanmasında etnik grup, yaş, cinsiyet ve diyabetin popülasyon prevalansı kaynaklı değişmelerin olacağı üzerine tartışmalar mevcuttur.

Treacy ve ark., 2001 yılında akut pankreatit üzerine yapılan çalışmalarında lipaz ve amilaz enzim seviyelerini değerlendirmişlerdir. ROC analizleri sonucunda 1. gün lipaz testi için en iyi kesim noktası 208 birim/l, duyarlılık %67, seçicilik %97, 3. gün lipaz testi için en iyi kesim noktası 216 birim/l, duyarlılık %55, seçicilik %84, 1. gün amilaz testi için en iyi kesim noktası 176 birim/l, duyarlılık %45, seçicilik %97, 3. gün amilaz testi için en iyi kesim noktası 126 birim/l, duyarlılık %35, seçicilik %92 olarak bulunmuştur. Lipaz ve amilaz enzimlerinin ROC eğrisi üzerinden performansları AUC değerleri ve güven aralıkları verilmeksizin sadece ROC eğrileri ile değerlendirilmiş olup lipaz testinin daha etkin ayırım gücüne sahip olduğu ifade edilmiştir (5).

Corsetti ve ark., 1993 yılında pankreatit üzerine yaptıkları çalışmalarında lipaz ve amilaz testi için sırasıyla ROC analizi sonuçlarına göre sırasıyla en iyi kesim noktası 196,5 birim/l ve 72,5 birim/l, duyarlılık %98,1 ve %72,6, seçicilik ise %96,3 ve %83,5 olarak bulunmuştur. Lipaz ve amilaz enzimlerinin sadece ROC eğrileri üzerinden performansları değerlendirilmiş olup lipaz testinin daha etkin ayırım gücüne sahip olduğu ifade edilmiştir (6).

Çalışmamızda lipaz testi için ROC analizi sonuçlarına göre en iyi kesim noktası 82 birim/l, duyarlılık %65,91, seçicilik %84,48, LR^+ 4,25, LR^- 0,4, PPV 0,74, NPV 0,76 olarak bulunmuştur. Lipaz testinden elde edilen AUC 0,751 değerinin (0,699-0,999) GA ile bulunup bunun kabul edilebilir bir ayırım gücüne sahip olduğu gözlenmiştir ($p<0,001$). Amilaz testi için ROC analizi sonuçlarına göre en iyi kesim noktası 90 birim/l, duyarlılık %63,64, seçicilik %72,41, LR^+ 2,31, LR^- 0,5, PPV 0,62, NPV 0,71 olarak bulunmuştur. Amilaz testinden elde edilen AUC 0,67 değerinin (0,614-0,722) güven aralığı ile bulunup bunun zayıf bir ayırım gücüne sahip olduğu gözlenmiştir ($p<0,001$). Yapılan çalışmaların var olan testlerine ait ROC analizi ölçütlerine bakıldığı zaman bazı ölçütlerin eksik olduğu görülmektedir. Var olan ölçütler ile çalışmamızı karşılaştıracak olursak verdiğimiz çalışmalarda lipaz testinin ayırım gücünün amilaz testine göre daha yüksek olduğu kanısı tutarlılık göstermektedir. Var olan farklılıkların ise yapılan ölçüm günlerinin kesin olmamakla birlikte farklı günler olmasından kaynaklanabileceği düşünülmektedir.

6. SONUÇ VE ÖNERİLER

Bu bölümde belirtilen klinik vakalar üzerinden uygulaması yapılan ROC analizi neticesinde ulaşılan bulguların yorumlaması yapılmış ve bu konu üzerinden yapılacak ileri ki çalışmalar için önerilerde bulunulmuştur.

6.1. Sonuçlar

Araştırmamızın çıkış noktası olan ROC analizinin, tanı testlerinin eşik değerlerinin belirlenmesinde ve elde edilen ölçütler neticesinde oluşan ROC eğrisi yardımıyla, ele alınan testin ayırım gücünün hesaplanmasında kullanıldığı çalışmanın başında ifade edilmiştir. Belirlenen klinik vakalar üzerinde analizimizin uygulaması yapılmış olup aşağıdaki sonuçlara ulaşılmıştır.

Diyabet klinik vakası için, HbA1c, glukoz, üre ve kreatinin testleri olmak üzere farklı tanı testinin sonuçları ele alınmıştır.

HbA1c tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 6,2 olarak belirlenmiştir. HbA1c testi için 6,2 değerine karşılık gelen duyarlılık değeri 69,51 ve seçicilik değeri 67,25 olarak bulunmuştur. Bulunan ihtimaller neticesinde HbA1c testinin diyabet olanları bulmakta daha başarılı olduğu ifade edilebilir. HbA1c testinden elde edilen AUC değeri ise 0,746 bulunmuş olup, kabul edilebilir bir ayırım gücüne sahiptir denilebilir ($p<0,001$).

Glukoz tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 116 olarak belirlenmiştir. Glukoz testi için 116 değerine karşılık gelen duyarlılık değeri 67,07 ve seçicilik değeri 72,49 olarak bulunmuştur. Bulunan ihtimaller sonucunda Glukoz testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir. Glukoz testinden elde edilen AUC değeri ise 0,726 bulunmuş olup, kabul edilebilir bir ayırım gücüne sahiptir denilebilir ($p<0,001$).

Üre tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 29 olarak belirlenmiştir. Üre testi için 29 değerine karşılık gelen duyarlılık değeri 41,46 ve seçicilik değeri 65,94 olarak bulunmuştur. Bulunan ihtimaller sonucunda üre testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir. Üre testinden elde edilen AUC değeri ise 0,53 bulunmuş olup, zayıf bir ayırım gücüne sahiptir denilebilir ($p=0,414$).

Kreatinin tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 0,8 olarak belirlenmiştir. Kreatinin testi için 0,8 değerine karşılık gelen duyarlılık değeri 74,39 ve seçicilik değeri 31,88 olarak bulunmuştur. Bulunan ihtimaller sonucunda Kreatinin testinin diyabet olanları bulmakta daha başarılı olduğu ifade edilebilir. Kreatinin testinden elde edilen AUC değeri ise 0,514 bulunmuş olup, zayıf bir ayırım gücüne sahiptir denilebilir ($p=0,669$).

Genel olarak, MedCalc'dan elde edilen AUC kestirimleri ile LABROC parametrik olmayan ve yarı parametrik yaklaşımlara göre elde edilen AUC kestirimleri arasında büyük farklılıklar olmadığı belirlenmiştir. AUC kestirimleri, standart hatalar ve güven aralıkları arasında çok küçük farklılıklar bulunmuştur. En yüksek farklılık Kreatinin testinin AUC kestiriminde görülmüş olup, önemli derecede etkili bir farklılık olmadığı belirtilebilir.

Diyabet tanısı için kullanılan testlerde, ROC eğrisi kestirimine göre 0,746 ile en yüksek AUC değerine ve 0,030 ile en düşük standart hataya sahip olan tanı testi HbA1c testidir. Dolayısıyla diyabet tanısında belirtilen kestirime göre en yüksek ayırım gücüne sahip olan testin HbA1c testi olduğunu söylemek mümkündür.

Akut pankreatit klinik vakası için, lipaz ve amilaz testleri olmak üzere farklı tanı testinin sonuçları ele alınmıştır.

Lipaz tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 82 olarak belirlenmiştir. Lipaz testi için 82 değerine karşılık gelen duyarlılık değeri 65,91 ve seçicilik değeri 84,48 olarak bulunmuştur. Bulunan ihtimaller neticesinde lipaz

testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir. Lipaz testinden elde edilen AUC değeri ise 0,751 bulunmuş olup, kabul edilebilir bir ayırım gücüne sahiptir denilebilir ($p<0,001$).

Amilaz tanı testinde duyarlılık ve seçicilik ölçüt değerinin farklarının ve ayırım gücünün en yüksek olduğu noktaya karşılık gelen eşik değeri 90 olarak belirlenmiştir. Amilaz testi için 90 değerine karşılık gelen duyarlılık değeri 63,64 ve seçicilik değeri 72,41 olup olarak bulunmuştur. Bulunan ihtimaller neticesinde Amilaz testinin diyabet olmayanları bulmakta daha başarılı olduğu ifade edilebilir. Amilaz testinden elde edilen AUC değeri ise 0,67 bulunmuş olup, p değeri ayırım gücü bakımından anlamlı bulunsa da zayıf bir ayırım gücüne sahiptir denilebilir ($p<0,001$).

Genel olarak, MedCalc'dan elde edilen AUC kestirimleri ile LABROC parametrik olmayan ve yarı parametrik yaklaşımlara göre elde edilen AUC kestirimleri arasında büyük farklılıklar olmadığı belirlenmiştir. AUC kestirimleri, standart hatalar ve güven aralıkları arasında çok küçük farklılıklar bulunmuştur.

Akut pankreatit tanısı için kullanılan testlerde, ROC eğrisi kestirimine göre 0,751 ile en yüksek AUC değerine ve 0,03 ile en düşük standart hataya sahip olan tanı testi lipaz testidir. Dolayısıyla akut pankreatit tanısında belirtilen kestirimlere göre en yüksek ayırım gücüne sahip olan testin lipaz testi olduğunu söylemek mümkündür.

6.2. Öneriler

Çalışmamızın sonuçlarından hareketle öncelikli olarak daha sonra yapılacak benzer çalışmalarda örneklem büyüklüğünün mümkün oldukça daha fazla olması tavsiye edilir. Bu hem kesinlik (accuracy) değerini arttıracak ve hem de kesinlik sınırını daha fazla arttırmaya neden olacaktır. Bunun yanında bilinen bir gerçek şu ki örneklem büyüklüğü arttıkça testin normal dağılım eğilimi artacaktır.

Ele alınan klinik vakada cinsiyet faktörünün önemli bir yere sahip olabileceği düşüncesiyle cinsiyet bazında analizler ve değerlendirmeler yapılabilir. Yukarıdaki

bulunan eřik deęerlerinin cinsiyet temelinde ne gibi farklılıkların veya benzerliklerin olabileceęi test tanım tahmini yapılabilir. ROC analizi bu testler için uygulanabilir.

Aynı Őekilde, yař faktörünün önemli olduęu vakalarda örneklemler yař gruplarına ayrılarak deęerlendirmeleri ayrıntılı bir Őekilde yapılabilir ve ROC analizi tanı ön testleri için uygulanabilir.



ÖZET

Klinik Vaka Tanısında Receiver Operating Characteristic (ROC) Analizi Tekniđi

Sađlık alanında klinisyenler ve arařtırmacılar eksik bilgi ve belirsiz sonuçlara rađmen sađlam kararlar vermek için tanı testlerine başvururlar. Bu testler tıbbi karar oluşumunda kullanılmadan önce performans deđerlendirmelerinin yapılması ve hastalık durumunu en yüksek ve dođru bir sınıflama yapılarak optimum eşik deđerinin bulunması gerekir. Bu deđerlendirmelerin ve işlemlerin yapılması açısından Receiver Operating Characteristic (ROC) analizi önemli bir istatistiksel tekniktir.

Bu çalışmada tanı deđerlendirme ölçütleri ile ROC analizini tanıtmak, kullanımı hakkında bilgi vermek ve anlatılan, verilen bilgilerin klinik vakalar üzerinden farklı istatistiksel programlar ile bir uygulamasının yapılması amaçlanmıştır. Bu anlatımlarla birlikte, tıbbi arařtırmacıların etkili bir şekilde kullanmasına yardımcı olmak için ROC analizinin temel kavramsal çerçevesini ve yorumunu sađlamaktır.

Uygulamada diyabet ve akut pankreatit olmak üzere iki farklı klinik vaka ele alınmıştır. Diyabet vakası için, şeker yükleme testi (OGTT) altın standart ölçütü olmak üzere HbA1c, glukoz, üre ve kreatinin testlerinin ayırım güçleri deđerlendirilmiştir. HbA1c testi, diyabet tanısı için kullanılan testlerde, ROC eğrisi kestirimine göre 0,746 ile en yüksek AUC deđerine ve 0,03 ile en düşük standart hataya sahip olan tanı testi olarak bulunmuştur ($p<0,001$). Akut pankreatit vakası için, bilgisayarlı tomografi (BT) altın standart ölçütü olmak üzere lipaz ve amilaz testlerinin ayırım güçleri deđerlendirilmiştir. Lipaz testi, akut pankreatit tanısı için kullanılan testlerde, ROC eğrisi kestirimine göre 0,751 ile en yüksek AUC deđerine ve 0,03 ile en düşük standart hataya sahip olan tanı testi olarak bulunmuştur ($p<0,001$).

Anahtar Kelimeler: ROC analizi, klinik vaka, tanı testleri

ABSTRACT

Receiver Operating Characteristic (ROC) Curve Analysis in Diagnosis of Clinical Cases

Clinicians and researchers in the field of health care make sound decisions, despite the incomplete information and uncertain results. Before these medical decisions, some performance evaluations must be made and the optimal threshold value should be calculated for an accurate classification for the disease. Receiver Operating Characteristic (ROC) analysis used as an important tool for statistical calculations for making pre-evaluations. It is moderately implemented in the healthcare domain to determine how accurately a patient's diagnostic test can be distinguished whether he or she has a disease or free from it.

In this study, it was aimed to introduce the diagnostic evaluation criteria and ROC analysis to give information about its usage, and to make the application of the information given through clinical cases with different statistical programs. We, thus, to provide the fundamental conceptual framework and interpretation of ROC analysis help medical researchers to use it effectively.

For ROC analysis, we investigated two different clinical cases including diabetes and acute pancreatitis. For diabetes, the different strengths of HbA1c, glucose, urea and creatinine tests were evaluated and examined. In the case of HbA1c test, which is used for diagnosis of diabetes, our ROC curve estimation showed that the highest Area under Curve (AUC) value was 0,746 with the lowest standard error 0,03 ($p < 0.001$). It is considered a fair discrimination. For acute pancreatitis, the discriminatory powers of the lipase and amylase assays were evaluated via the gold standard of computer tomography. In the event of lipase test, which is for the diagnosis of acute pancreatitis, our ROC curve estimation showed that the highest AUC value was 0,751 with the lowest standard error 0,03 ($p < 0.001$). It is also expressed as a fair discrimination.

Key Words: ROC analysis, clinical case, diagnostic tests

KAYNAKLAR

1. Baykul Y. Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması. Ankara: Pegem, 2010.
2. Krzanowski JW, David JH. ROC Curves for Continuous Data. Chapman and Hall/CRC, 2009.
3. Dirican A. Tanı testi performansının değerlendirilmesi ve kıyaslanması. Cerrahpaşa Tıp Dergisi 2001; 32: 25-30.
4. Lusted LB. Signal detectability and medical decision making. Science 1971; Vol.171, 3977:1217-1219.
5. Green DM, Swets JA. Signal detection theory and psychophysics. Huntington NY Krieger 1974; 401-404.
6. Goodenough DJ, Rossmann K, Lusted LB. Radiographic applications of signal detection theory. Radiology 1972; 105:199-200.
7. Egan JP. Signal Detection Theory and ROC Analysis. Academic Press 1975; New York.
8. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. American Journal of Roentgenology 2005; Vol.184-1.
9. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. Clinical Chemistry 1993; 39: 561-77.
10. Metz CE. Basic principles of ROC analysis. Seminars in Nuclear Medicine 1978; 8:283-98.
11. Hanley JA, McNeil, BJ. The meaning and use of the area under an ROC curve. Radiology 1982; 143: 29-36.
12. Hanley JA ve McNeil BJ. A method of comparing the areas under receiver operating characteristic (ROC) curves derived from the same Cases. Radiology 1983; Vol.148, 3: 839-843.
13. Kim MJ, Lim JS, Oh YT. Preoperative MRI of rectal cancer with and without rectal water filling: an intraindividual comparison. AJR Am J Roentgenol 2004; 182:1469-1476.

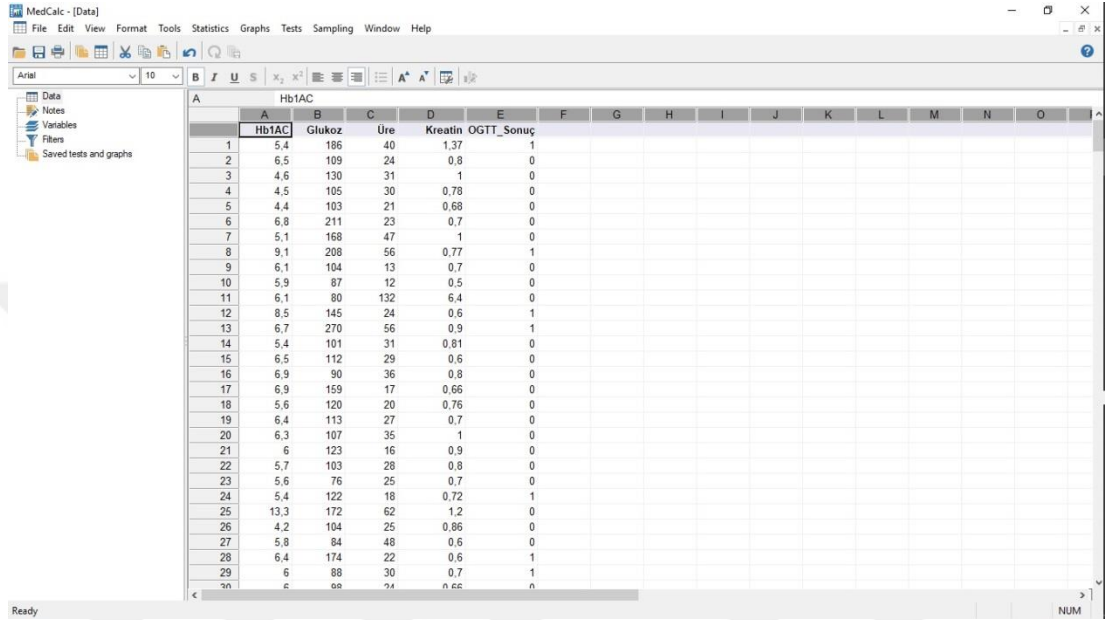
14. Osada H, Kaku K, Masuda K, Litsuka Y, Seki K, Sekiya S. Quantitative and qualitative evaluations of fetal lung with MR imaging. *Radiology* 2004; 231(3):887-92.
15. Zheng B, Leader JK, Abrams G. Computer-aided detection schemes: The effect of limiting the number of cued regions in each case. *American Journal of Roentgenology* 2004; 182(3):579-83.
16. Nicoll D, Lu CM, Pignone M, McPhee SJ. *Pocket Guide to Diagnostic Tests. Medicine* 6th Ed. 2013, ISBN-10: 0071310657.
17. Sox HC, Higgins MC, Owens DK. *Medical Decision Making*. John Wiley 2013.
18. Kassirer P. The threshold approach to clinical decision making. *J. Med.* 1980; 302:1109-1117.
19. Dawson B, Trapp R. *Basic and Clinical Biostatistic*. 4th Ed. 2004, ISBN:007141018X
20. Obuchowski NA. Computing sample size for receiver operating characteristic studies. *Investigative Radiology* 1994; 29: 238-43.
21. Shapiro DE. The interpretation of diagnostic tests. *Statistical Methods in Medical Research* 1999; 8: 113-34.
22. Zhou XH, McClish DK, Obuchowski NA. *Statistical Methods in Diagnostic Medicine*. John Wiley 2002, New York.
23. Pepe MK. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press 2003, ISBN 978-0-19-850984-4.
24. Shapiro DE. The interpretation of diagnostic tests. *Statistical Methods in Medical Research* 1999; 8: 113-34.
25. Boyd JC. Mathematical tools for demonstrating the clinical usefulness of biochemical markers. *Scandinavian Journal of Clinical and Laboratory Investigation* 1997; 227: 46-63.
26. Boyko EJ. Ruling out or ruling in disease with the most sensitive or specific diagnostic test: Short cut or wrong turn? *Medical Decision Making* 1994; 14:175-9.
27. Fagan TJ. Letter: Nomogram for Bayes theorem. *The New England journal of Medicine* 1975; 5: 257.

28. Köksal B. Regresyon analizinde ROC eğrisi kestirimi ile model seçimi. Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, 2011.
29. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics* 2005; 38: 404-415.
30. Gail MH, Green SB. A generalization of the one-sided twosample Kolmogorov-Smirnov statistic for evaluating diagnostic tests. *Biometrics* 1976; 32:561-70.
31. ADA. Standarts of medical care in diabetes. *Diabetes Care* 2008; 31: 12-20.
32. WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia: report of a WHO/IDF consultation. Geneva, Switzerland: WHO 2006; 1-50.
33. Andersen DK, Biliary TR, Dunn DL, Hunter JG. Schwartz's principles of surgery. Textbook 2015; 1341-1422.
34. Akarsu M. Hastane verilerinden hareketle diyabet teşhisinde kullanılan biyokimyasal testlerin tanısal değerlendirilmesi. Uzmanlık Tezi, Haydarpaşa Numune Hastanesi, 2008.
35. Colagiuri S, Cameron A, Hussain Z, Shaw J, Zimmet P. Screening for type 2 diabetes and impaired glucose metabolism: the Australian experience. *Diabetes Care* 2004; 27: 367–371.
36. Mannucci E, Ognibene A, Sposato I, Brogi M, Gallori G, Bardini G et al. Fasting plasma glucose and glycated haemoglobin in the screening of diabetes and impaired glucose tolerance. *Acta Diabetol* 2003; 40: 181–186.
37. Tavintharan S, Chew LSW, Heng DMK. A rational alternative for the diagnosis of diabetes mellitus in high risk individuals. *Ann Acad Med Singapore* 2000; 29: 213–218.
38. Treacy J, Williams A, Bais R, Willson K et al. Evaluation of amylase and lipase in the diagnosis of acute pancreatitis. *Anz J. Surg.* 2001; 71:577-582.
39. Corsetti JP, Cox C, Shulz TJ, Avan DA. Combined serum amylase and lipase determinations for diagnosis of suspected acute pancreatitis. *Clin. Chem.* 1993; 39/12: 2495-2499.

EKLER

Ek 1. MedCalc ve LABROC Ekran Görüntüleri

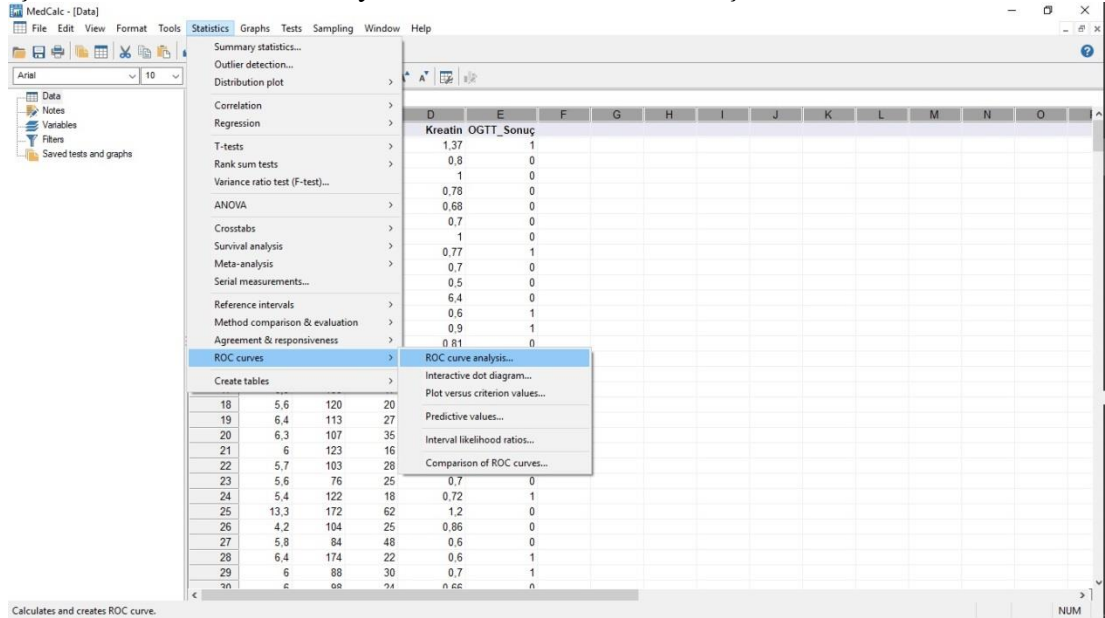
İşlem 1. Verilerin MedCalc programına aktarılması



The screenshot shows the MedCalc software interface with a data table. The table has columns for Hb1AC, Glukoz, Üre, Kreatin, and OGTT_Sonuç. The data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Hb1AC	Glukoz	Üre	Kreatin	OGTT_Sonuç										
1	5,4	186	40	1,37	1										
2	6,5	109	24	0,8	0										
3	4,6	130	31	1	0										
4	4,5	105	30	0,78	0										
5	4,4	103	21	0,68	0										
6	6,8	211	23	0,7	0										
7	5,1	168	47	1	0										
8	9,1	208	56	0,77	1										
9	6,1	104	13	0,7	0										
10	5,9	87	12	0,5	0										
11	6,1	80	132	6,4	0										
12	8,5	145	24	0,6	1										
13	6,7	270	56	0,9	1										
14	5,4	101	31	0,81	0										
15	6,5	112	29	0,6	0										
16	6,9	90	36	0,8	0										
17	6,9	159	17	0,66	0										
18	5,6	120	20	0,76	0										
19	6,4	113	27	0,7	0										
20	6,3	107	35	1	0										
21	6	123	16	0,9	0										
22	5,7	103	28	0,8	0										
23	5,6	76	25	0,7	0										
24	5,4	122	18	0,72	1										
25	13,3	172	62	1,2	0										
26	4,2	104	25	0,66	0										
27	5,8	94	48	0,6	0										
28	6,4	174	22	0,6	1										
29	6	88	30	0,7	1										
30	6	88	24	0,66	0										

İşlem 2. ROC analiz fonksiyonunun analiz menüsünden seçimi



The screenshot shows the MedCalc software interface with the Statistics menu open. The ROC curves option is selected, and the ROC curve analysis... option is highlighted. The data table is visible in the background.

	D	E	F	G	H	I	J	K	L	M	N	O
	Kreatin	OGTT_Sonuç										
1	1,37	1										
2	0,8	0										
3	1	0										
4	0,78	0										
5	0,68	0										
6	0,7	0										
7	1	0										
8	0,77	1										
9	0,7	0										
10	0,5	0										
11	6,4	0										
12	0,6	1										
13	0,9	1										
14	0,81	0										
15	0,6	0										
16	0,8	0										
17	0,66	0										
18	0,76	0										
19	0,7	0										
20	1	0										
21	0,9	0										
22	0,8	0										
23	0,7	0										
24	0,72	1										
25	1,2	0										
26	0,66	0										
27	0,6	0										
28	0,6	1										
29	0,7	1										
30	0,66	0										

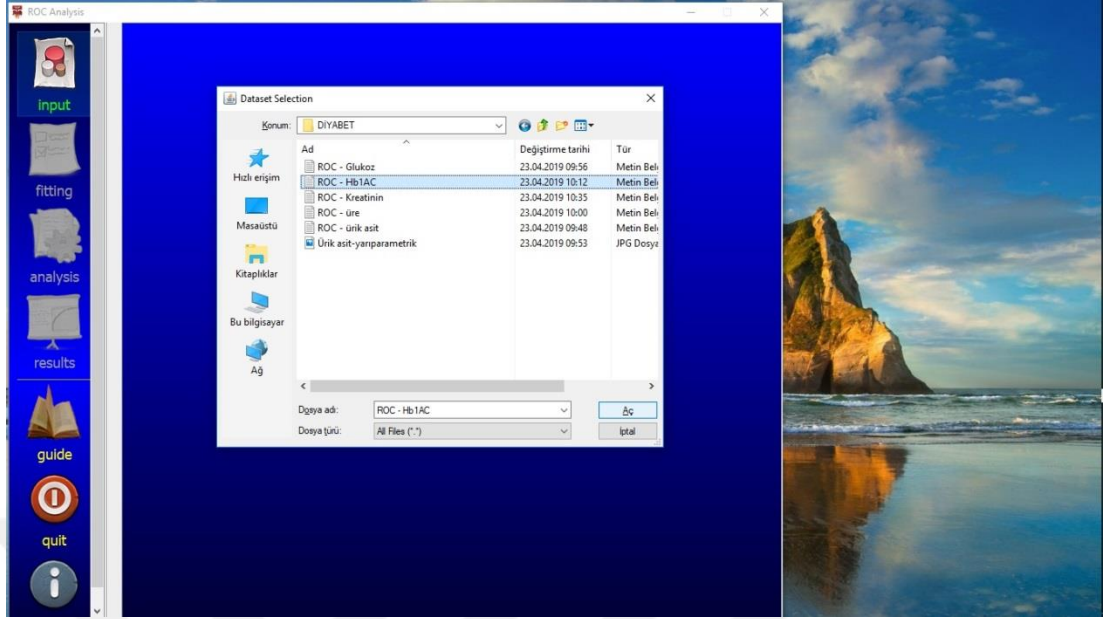
İşlem 3. ROC analizi için değişkenlerin analiz menüsünde yerleştirilmesi

Complete dialog box (press F1 for Help).

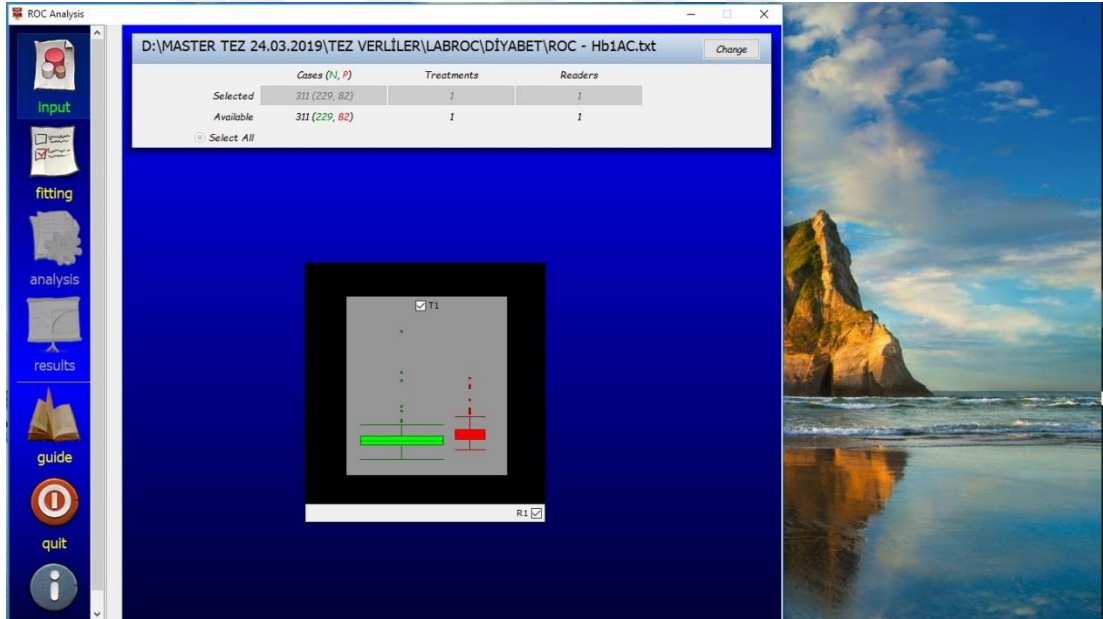
İşlem 4. ROC analiz sonucu

Variable	Hb1AC							
Classification variable	OGTT_Sonuç							
Sample size	311							
Positive group ^a	82 (26.37%)							
Negative group ^b	229 (73.63%)							
Disease prevalence (%)	unknown							
Area under the ROC curve (AUC)								
Area under the ROC curve (AUC)	0.746							
Standard Error ^a	0.0309							
95% Confidence interval ^b	0.694 to 0.793							
z statistic	7.950							
Significance level P (Area=0.5)	<0.0001							
^a DeLong et al., 1988								
^b Binomial exact								
Youden index								
Youden index J	0.3676							
Associated criterion	>6.2							
Sensitivity	69.51							
Specificity	67.25							
Criterion values and coordinates of the ROC curve [Hide]								
Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	95% CI	-LR	95% CI
>4	100.00	95.6 - 100.0	0.00	0.0 - 1.6	1.00	1.0 - 1.0		
>4	100.00	95.6 - 100.0	0.44	0.01 - 2.4	1.00	1.0 - 1.0	0.00	
>4.1	100.00	95.6 - 100.0	1.31	0.3 - 3.8	1.01	1.0 - 1.0	0.00	
>4.2	100.00	95.6 - 100.0	1.75	0.5 - 4.4	1.02	1.0 - 1.0	0.00	
>4.3	100.00	95.6 - 100.0	2.18	0.7 - 5.0	1.02	1.0 - 1.0	0.00	

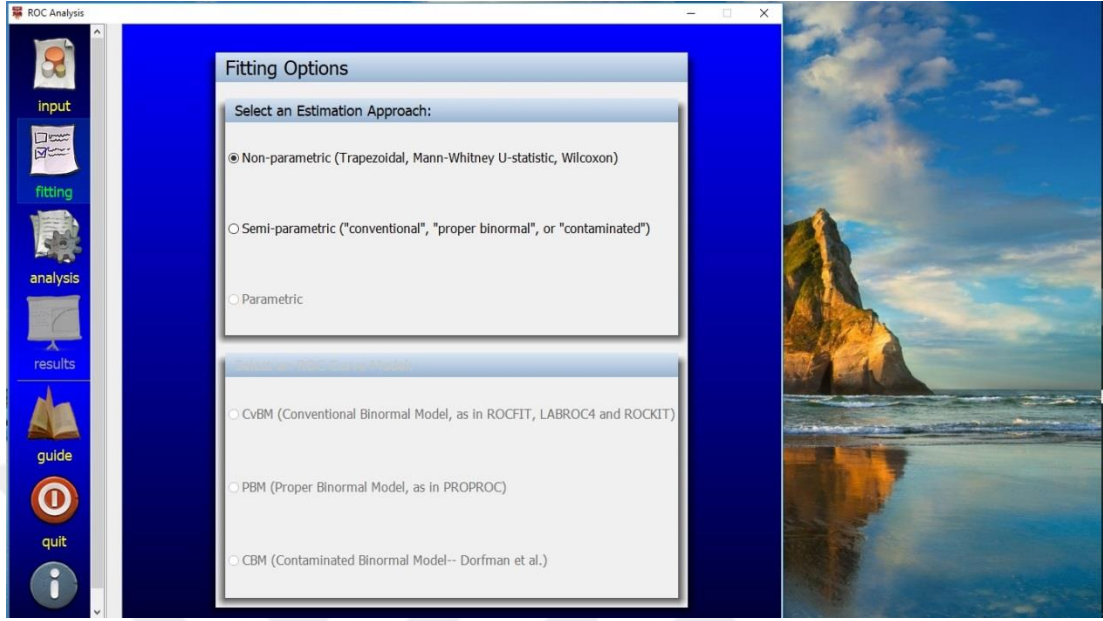
İşlem 5. LABROC programında verilerin atanması



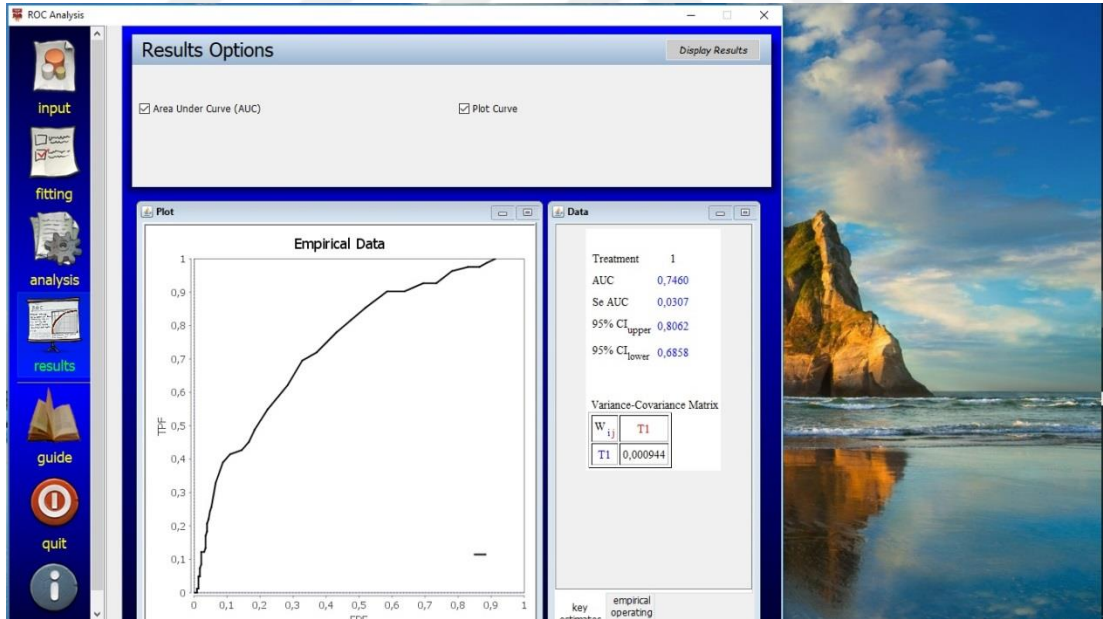
İşlem 6. LABROC programında karar matrisi sonucunun gösterimi



İşlem 7. LABROC programından yarı parametrik ve parametrik olmayan metotların seçimi



İşlem 8. LABROC programında ROC analizi sonucunun gösterimi



Ek 2. Etik Kurul İzni



T.C.
SÜLEYMAN DEMİREL ÜNİVERSİTESİ
Tıp Fakültesi Klinik Araştırmalar Etik Kurulu Başkanlığı

Sayı : 72867572.050.01.04- *55202*
Konu : Etik Kurul Kararı

10 -04- 2019

Sayın Dr. Öğr. Üyesi Osman GÜRDAL
Tıp Fakültesi
Biyoistatistik ve Tıbbi Bilişim A.D.

Sorumlu araştırmacı olduğunuz “Klinik Vaka Tanısında ROC Analizi Tekniği” isimli çalışmanızın kurulumuz tarafından uygun görüldüğüne ilişkin 05/04/2019 tarih ve 127 sayılı Süleyman Demirel Üniversitesi Tıp Fakültesi Klinik Araştırmalar Etik Kurulu Kararı yazımız ekinde gönderilmiştir.
Bilgilerinizi rica ederim.


Prof. Dr. Mekin SEZİK
Başkan

Eki : 1 Adet Etik Kurulu Kararı (2 Sayfa)

S.D.Ü. Tıp Fakültesi Dekanlığı Doğu Kampusu 32260 - ISPARTA
Tel : 0 (246) 2113704 Faks : 0 (246) 2371165
e-posta : tipetik@sdu.edu.tr İnternet Adresi : www.tip.sdu.edu.tr

Bilgi İçin : İ.Etem YETİŞEN
Bilgisayar İşletmeni
Tel : 0 (246) 2113704

Ek 3. Özgeçmiş

Kişisel Bilgiler

Ad / Soyad : Ebru DUMLUPINAR
Doğum Yeri : Gölhisar / BURDUR
Doğum Tarihi : 12.05.1989
Telefon No : 0554 940 2478
E- mail : edumlupinar@adiyaman.edu.tr

Eğitim Bilgileri

Lise : Gölhisar Y.D.A. Lise
Lisans : Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi
Matematik Bölümü (2008 - 2012)

İş Denevim Bilgileri

Görevi : Araştırma Görevlisi
Kurumu : Adıyaman Üniversitesi Tıp Fakültesi Temel Tıp Bilimleri Bölümü
Biyoistatistik ve Tıp Bilişimi Anabilim Dalı

Yabancı Diller

YDS / YÖKDİL : 65 (İngilizce)