

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**ZAMANA BAĞLI ROC ANALİZİ VE SAĞLIK ALANINDA
UYGULAMALARI**

Ceren EFE

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI
YÜKSEK LİSANS TEZİ

DANIŞMANI

Yrd. Doç. Dr. İlker ÜNAL

ADANA-2017

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**ZAMANA BAĞLI ROC ANALİZİ VE SAĞLIK ALANINDA
UYGULAMALARI**

Ceren EFE

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI
YÜKSEK LİSANS TEZİ

DANIŞMANI

Yrd. Doç. Dr. İlker ÜNAL

ADANA-2017

KABUL VE ONAY

Biyoistatistik Anabilim Dalı

Tezli Yüksek Lisans Programı Çerçevesinde yürütülmüş olan
“Zamana Bağlı ROC Analizi ve Sağlık Alanında Uygulamaları”
adlı çalışma, aşağıdaki jüri tarafından **Yüksek Lisans Tezi** olarak kabul edilmiştir.

Tarihi: 03/07/2017

TEZ SINAV JÜRİSİ

Yrd. Doç. Dr. İlker ÜNAL
Çukurova Üniversitesi
Başkan

Prof. Dr. Z. Nazan ALPARSLAN
Çukurova Üniversitesi
Üye

Doç. Dr. Gülhan ÖREKİCİ TEMEL
Mersin Üniversitesi
Üye

Yukarıdaki Tez, Yönetim Kurulunun / / tarih ve sayılı kararı ile kabul edilmiştir.

Prof.Dr. Behice DURGUN
Sağlık Bilimleri Enstitü Müdürü

T.C. ÇUKUROVA ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI
ETİK BEYANI

Çukurova Üniversitesi Bilimsel Araştırma ve Yayın Etiği Yönergesini okuduğumu ve anladığımı ve Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde ve ortaya çıkan sonuçlarda herhangi bir değişiklik yapmadığımı,
- Tez olarak sunduğum bu çalışmanın özgün olduğunu,

bildirir, aksi bir durumda bu konuda hakkımda yapılacak tüm yasal işlemleri ve aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 03/07/2017

Ceren EFE

Kayıtlı olunan Program : Biyoistatistik Anabilim Dalı
Tezin Konusu : Zamana Bağlı ROC Analizi ve Sağlık Alanında Uygulamaları

Tezin Türü : Yüksek Lisans: Doktora:

Danışmanın Adı-Soyadı : Yrd.Doç.Dr. İlker ÜNAL

Danışmanın İletişim Bilgileri

Telefon : 0 (322) 338 60 60 - 3094

E-Posta : ilkerun@cu.edu.tr

Öğrencinin İletişim Bilgileri

Telefon : 0 (554) 713 91 07

E-Posta : cefe@cu.edu.tr

Adresi : Çukurova Üniversitesi Tıp Fakültesi
Biyoistatistik ABD Sarıçam/Adana

**Bu belgenin Lisansüstü eğitim tezleri savunmaya alınmadan önce öğrenci tarafından doldurulup imzalanarak Enstitü Müdürlüğüne teslim edilmesi gerekmektedir.*

TEŐEKKÜR

Tezimin planlanmasında ve yürütülmesinde yardımlarını esirgemeyen tez danışmanım Yrd. Doç. Dr. İlker Ünal'a; eleştirileri, katkıları ve düzeltmeleri için hocalarım Prof. Dr. Z. Nazan Alparslan, Prof. Dr. H. Refik Burgut, Prof. Dr. Gülşah Seydaođlu ve Doç. Dr. Yaşar Sertdemir'e; çalışmada kullanılan veri setlerinden birini sağlayan Prof. Dr. Semra Paydaş ve çalışma grubuna; bölüm arkadaşlarıma; akademik ve manevi desteklerinden ötürü sevgili annem Prof. Dr. Lale Efe ve sevgili babam Prof. Dr. Ercan Efe'ye sonsuz teşekkürlerimi sunarım.

Ceren EFE



İÇİNDEKİLER

| | |
|-----------------------------------------------------------------------------------------------------------------------|------|
| KABUL VE ONAY | i |
| ETİK BEYANI | ii |
| TEŞEKKÜR..... | iii |
| İÇİNDEKİLER | iv |
| ŞEKİLLER DİZİNİ | vi |
| ÇİZELGELER DİZİNİ | viii |
| SİMGELER VE KISALTMALAR DİZİNİ | x |
| ÖZET | xi |
| ABSTRACT..... | xii |
| 1. GİRİŞ..... | 1 |
| 2. GENEL BİLGİLER..... | 6 |
| 2.1. Tanı Testleri..... | 6 |
| 2.2. Tanı Testlerinin Değerlendirilmesinde Kullanılan Doğruluk Ölçütleri | 6 |
| 2.2.1. Nitel Ölçümlü Testlerde Doğruluk | 7 |
| 2.2.2. Nicel Ölçümlü Testlerde Doğruluk..... | 9 |
| 2.2.2.1. ROC Analizi..... | 9 |
| 2.3. Zamana-Bağlı ROC Analizi | 12 |
| 2.3.1. Boylamsal Verilerde Zamana-Bağlı ROC Analizi | 13 |
| 2.3.1.1. Verinin Modellenmesi..... | 16 |
| 2.3.1.2. ROC Eğrisinin Modellenmesi | 18 |
| 2.3.2. Sağkalım Verilerinde Zamana-Bağlı ROC Analizi | 20 |
| 2.3.2.1. Kaplan-Meier Tahmin Edicisi:..... | 22 |
| 2.3.2.2. İkili Dağılımın En Yakın Komşu Tahmin Edicisi (NNE):..... | 24 |
| 3. GEREÇ VE YÖNTEM..... | 26 |
| 3.1. Kolon Kanseri Verisi..... | 26 |
| 3.2. Melano Verisi | 27 |
| 3.3. Paquid Verisi | 27 |
| 3.4. Hodgkin Lenfoma Verisi | 27 |
| 4. BULGULAR | 29 |
| 4.1. Klasik ROC Yöntemi ile Zamana-Bağlı ROC Analizi Yöntemi olan Kaplan-Meier Kestirimlerinin Karşılaştırılması..... | 29 |

| | |
|-------------------------------------------------------------------------------------------|----|
| 4.1.1. Kolon Kanseri Verisinden Elde Edilen Bulgular..... | 29 |
| 4.1.2. Melano Verisinden Elde Edilen Bulgular | 33 |
| 4.1.3. Paquid Verisinden Elde Edilen Bulgular | 37 |
| 4.1.3.1. DSST Belirteç Olarak Alındığında | 38 |
| 4.1.3.2. MMSE Belirteç Olarak Alındığında | 41 |
| 4.1.4. Hodgkin Lenfoma Verisinden Elde Edilen Bulgular..... | 44 |
| 4.1.4.1. IPS7 Skoru Belirteç Olarak Alındığında..... | 45 |
| 4.1.4.2. IPS3 Skoru Belirteç Olarak Alındığında..... | 48 |
| 4.2. Kaplan-Meier (KM) ile En Yakın Komşu (NNE) Kestirimlerinin Karşılaştırılması..... | 51 |
| 4.2.1. Kolon Kanseri Verisinden Elde Edilen Bulgular..... | 51 |
| 4.2.2. Melano Verisinden Elde Edilen Bulgular | 54 |
| 4.2.3. Paquid Verisinden Elde Edilen Bulgular | 57 |
| 4.2.3.1. DSST Belirteç Olarak Alındığında | 57 |
| 4.2.3.2. MMSE Belirteç Olarak Alındığında | 60 |
| 4.2.3.3. DSST ve MMSE Belirteçlerinin Karşılaştırılması..... | 63 |
| 4.2.4. Hodgkin Lenfoma Verisinden Elde Edilen Bulgular..... | 64 |
| 4.2.4.1. IPS7 Skoru Belirteç Olarak Alındığında..... | 64 |
| 4.2.4.2. IPS3 Skoru Belirteç Olarak Alındığında..... | 67 |
| 5. TARTIŞMA..... | 70 |
| 6. SONUÇ VE ÖNERİLER | 75 |
| KAYNAKLAR | 76 |
| EKLER..... | 79 |
| ÖZGEÇMİŞ | 84 |

ŞEKİLLER DİZİNİ

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Şekil 2.1. ROC Eğrisi | 11 |
| Şekil 4.1. Kolon kanseri veri seti için pozitif nodül sayısı belirtecine ait klasik ROC eğrisi | 30 |
| Şekil 4.2. Kolon kanseri veri seti için pozitif nodül sayısı belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 32 |
| Şekil 4.3. Melano veri seti için tümör kalınlığı belirtecine ait klasik ROC eğrisi | 34 |
| Şekil 4.4. Melano veri seti için tümör kalınlığı belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 36 |
| Şekil 4.5. Paquid veri seti için DSST belirtecine ait klasik ROC eğrisi | 38 |
| Şekil 4.6. Paquid veri seti için DSST belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 40 |
| Şekil 4.7. Paquid veri seti için MMSE belirtecine ait klasik ROC eğrisi | 41 |
| Şekil 4.8. Paquid veri seti için MMSE belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 43 |
| Şekil 4.9. Hodgkin Lenfoma veri seti için IPS7 skoruna ait klasik ROC eğrisi | 45 |
| Şekil 4.10. Hodgkin Lenfoma veri seti için IPS7 skoruna ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 47 |
| Şekil 4.11. Hodgkin Lenfoma veri seti için IPS3 skoruna ait klasik ROC eğrisi | 48 |
| Şekil 4.12. Hodgkin Lenfoma veri seti için IPS3 skoruna ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği | 50 |
| Şekil 4.13. Farklı zamanlardaki pozitif nodül sayısı için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 52 |
| Şekil 4.14. Kolon kanseri verisi üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin karşılaştırılması değişik zaman noktalarında karşılaştırılması | 53 |

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Şekil 4.15. Farklı zamanlardaki tümör kalınlığı ölçümleri için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 55 |
| Şekil 4.16 Melanoma verisi üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin karşılaştırılması değişik zaman noktalarında karşılaştırılması | 56 |
| Şekil 4.17. Farklı zamanlardaki DSST skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 58 |
| Şekil 4.18 Paquid verisinde bulunan DSST belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması | 59 |
| Şekil 4.19. Farklı zamanlardaki MMSE skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 61 |
| Şekil 4.20. Paquid verisinde bulunan MMSE belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması | 62 |
| Şekil 4.21. Paquid veri seti için DSST ve MMSE belirteçlerinin KM ve NNE yöntemleri ile farklı zamanlarda karşılaştırılması | 63 |
| Şekil 4.22. Farklı zamanlardaki IPS7 skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 65 |
| Şekil 4.23. Hodgkin Lenfoma verisinde bulunan IPS7 belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması | 66 |
| Şekil 4.24. Farklı zamanlardaki IPS3 skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği | 68 |
| Şekil 4.25. Hodgkin Lenfoma verisinde bulunan IPS3 belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması | 69 |

ÇİZELGELER DİZİNİ

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Çizelge 2.1. Tanı Testlerinin Performansını Değerlendirmede Kullanılan 2x2 Karar Matrisi ^[25] | 7 |
| Çizelge 4.1. Kolon kanseri verisinin örnek görüntüsü | 29 |
| Çizelge 4.2. Kolon kanseri verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları | 29 |
| Çizelge 4.3. Kolon kanseri veri seti için 3 yıl-6 ay zaman aralığında elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 31 |
| Çizelge 4.4. Melano verisinin örnek görüntüsü | 33 |
| Çizelge 4.5. Melano verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları | 33 |
| Çizelge 4.6. Melano veri seti için 6 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 35 |
| Çizelge 4.7. Paquid verisinin örnek görüntüsü | 37 |
| Çizelge 4.8. Paquid verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları | 37 |
| Çizelge 4.9. Paquid veri seti için DSST belirtecine 12 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 39 |
| Çizelge 4.10. Paquid veri seti için MMSE belirtecine ait 12 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 42 |
| Çizelge 4.11. Hodgkin Lenfoma verisinin örnek görüntüsü | 44 |
| Çizelge 4.12. Hodgkin Lenfoma verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları | 44 |
| Çizelge 4.13. Hodgkin Lenfoma veri seti için IPS7 skoruna ait 6 yıl boyunca her 6 ay için elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 46 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Çizelge 4.14. Hodgkin Lenfoma veri seti için IPS3 skoruna ait 6 yıl boyunca her 6 ay için elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları | 49 |
| Çizelge 4.15. Kolon kanseri veri seti için 3 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 51 |
| Çizelge 4.16. Melano veri seti için 6-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 54 |
| Çizelge 4.17. Paquid veri setindeki DSST skoru için 12-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 57 |
| Çizelge 4.18. Paquid veri setindeki MMSE skoru için 12-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 60 |
| Çizelge 4.19. Hodgkin Lenfoma veri setindeki IPS7 skoru için 6 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 64 |
| Çizelge 4.20. Hodgkin Lenfoma veri setindeki IPS3 skoru için 6 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları | 67 |

SİMGELER VE KISALTMALAR DİZİNİ

| | |
|-------|------------------------------------------------------------------|
| ROC | Receiver Operating Characteristic (Alıcı İşletim Karakteristiği) |
| AUC | Area Under the Curve (Eğri Altında Kalan Alan) |
| Duy | Duyarlılık |
| Seç | Seçicilik |
| DPO | Doğru Pozitiflik Oranı |
| HPO | Hatalı Pozitiflik Oranı |
| KM | Kaplan-Meier |
| NNE | Nearest Neighbor Estimator (En Yakın Komşu Kestirimi) |
| DSST | Digit Symbol Substitution Score Test |
| MMSE | Mini Mental State Examination |
| NSABP | National Surgical Adjuvant Breast and Bowel Project |
| HL | Hodgkin Lenfoma |
| IPS | International Prognostic Score |
| EAKA | Eğri Altında Kalan Alan |
| GA | Güven Aralığı |

ÖZET

Zamana Bağlı ROC Analizi Ve Sağlık Alanında Uygulamaları

ROC analizi, hastalık tanısında kullanılan niceliksel tanı testlerinin veya biyolojik belirteçlerin sınıflama başarısını ölçmeye yarayan popüler bir yöntemdir. Ancak sağkalım verilerinde veya boylamsal verilerde olduğu gibi, ilgilenilen olay ile ölçüm yapılan belirteç arasında bir zamana bağlılık söz konusu ise, klasik ROC analizi belirtecin gerçek performansını tahmin edemeyebilir. Bu nedenle, sonuçları zamana bağlı olarak değişkenlik gösteren bir biyolojik belirtecin tanı performansını değerlendirirken zaman kavramını da hesaba katarak işlem yapan Zamana Bağlı ROC Analizi yöntemi geliştirilmiştir. Bu tez çalışmasında, ROC analizinin zamana bağlı durumlara uyarlanmış bu versiyonu incelenmiş ve çeşitli alanlardan elde edilmiş veriler üzerinde uygulamaları yapılmıştır.

Zamana bağlı ROC analizi yöntemlerinden Kaplan-Meier (KM) kestiriminin klasik ROC analizi ile karşılaştırması yapılmış, kullanılan tüm veri setlerinde zamanı göz önünde bulundurmanın klasik ROC'a göre daha başarılı olduğu saptanmıştır. Ayrıca KM yaklaşımı kullanılarak elde edilen, tüm zaman noktalarında yapılan ölçüm değerlerine ait ROC eğrileri altında kalan alanlar (EAKA) kendi aralarında karşılaştırılmıştır. Genel olarak, olayın gerçekleşmesine yaklaştıkça belirtecin performansının arttığı; özellikle 1 yıl veya 6 ay kala yapılan ölçümlerin daha eski ölçümlere göre daha belirleyici olduğu tespit edilmiştir.

KM kestirime alternatif olarak önerilmiş bir yaklaşım olan En Yakın Komşu kestirimi (Nearest Neighbor Estimator - NNE) ile yapılan uygulamalar sonucunda her iki yöntem karşılaştırılmıştır. Tüm veri setlerinde KM ve NNE yaklaşımları benzer sonuçlar vermiştir. KM için elde edilen ROC eğrilerinin EAKA değerleri genellikle NNE'den daha yüksek olsa da, sansürlü verilerde belirtecin değerlendirilmesinde NNE yaklaşımına başvurulmasının daha uygun olabileceği düşünülmektedir.

Bu tez çalışması ile, olay ve/veya belirteç üzerinde zamanın etkisinin bulunduğu durumlarda belirtecin daha doğru değerlendirilmesi adına klasik ROC analizi yerine zamana bağlı ROC analizi yöntemlerinin tercih edilmesinin daha doğru olacağı sonucuna varılmıştır.

Anahtar kelimeler: Zamana bağlı ROC, Tanı testleri, Biyolojik belirteç, Kaplan-Meier, En yakın komşu kestirimi

ABSTRACT

Time-Dependent ROC Analysis and Applications in the Field of Medicine

ROC analysis is a commonly used method for evaluating the accuracy of a continuous diagnostic test or a biomarker. However, when there is a time-dependency between the biomarker and the event of interest (death, disease, relapse etc.), using classical ROC analysis may not be able to estimate the true performance of the biomarker. For such cases, time-dependent ROC, an extended version of the standard ROC analysis, is developed. In this thesis, this modified ROC analysis adapted to the time dependent situations is investigated and applications are performed using datasets derived from several medical fields.

A comparison between classical ROC analysis and Kaplan-Meier (KM) estimator, which is a time-dependent ROC analysis method has been made on all datasets used in this study. From this comparison, it is concluded that time-dependent ROC method is superior to the standard ROC analysis. In addition, performance of biomarkers measured at each time-point are compared. In general, the closer to the event time, the higher performance is observed. Especially, biomarkers measured at last 12 or 6 months before the event are determined to be better at classification than the earlier measurements.

Nearest Neighbor Estimator (NNE), an alternative to KM estimator, is also applied on all datasets. Then, the findings obtained from these two approaches are compared. In all datasets, KM and NNE applications yielded very similar results. Although the KM-AUC values are mostly higher than the NNE-AUC values, it is more appropriate to use NNE method to evaluate the performance of a biomarker when time dependency exists in a censored data.

According to the results obtained from this study, when there is a time dependency between the biomarker and the event of interest, to measure a diagnostic performance of the biomarker accurately, the time-dependent ROC analysis is recommended instead of the classical ROC analysis.

Keywords: Time-dependent ROC, Diagnostic tests, Biomarker, Kaplan-Meier, Nearest Neighbor Estimator

1. GİRİŞ

Bir bireye doğru şekilde hastalık tanısı koymak, hasta kişileri sağlıklılarından mümkün olduğunca hatasız ayırt etmek sağlık alanında büyük ölçüde önem taşır. Bu ayrımı yüzde yüz kesinlikle yapan testlere altın standart test adı verilir. Ancak altın standart testlerin kullanımı her koşulda uygun olmayabilir. Bunun bir nedeni, kullanılacak olan altın standart testin bazen hastaya yararından çok zararının dokunabileceği ihtimalidir. Örneğin; amniyosentez, kromozom anomalilerinin tespitinde kesin sonuç verir ancak çok yüksek risk taşıdığından son çare olmadıkça tercih edilmemektedir. Aynı şekilde mamografi belli yaşın altındaki kadınlarda tarama amaçlı kullanılmamaktadır. Bir başka neden ise altın standart testlerin maliyetinin çok yüksek olması ve gerekmedikçe kullanılmasının doğru olmamasıdır. Bu nedenle altın standardın yerine, tanı testi adı verilen, ayırım başarısı yüzde yüz olmayan ancak daha düşük riskli ve maliyetli testler tercih edilmektedir.

Hasta olmayana tedavi uygulamak veya hasta olanı tedavisiz bırakmak yanlış tanı sonucunda karşılaşılan ciddi bir problemdir. Başarılı bir tanı testi ile hem gereksiz, hem de eksik tedavi uygulamaları minimuma indirilmiş olur. Hasta kişiler için izlenecek yolun ve uygulanacak tedavinin belirlenmesi, dolayısıyla da hastalığın etkili bir şekilde iyileştirilmesi için ilk adım doğru tanıdır. Tanı testlerinin ne ölçüde doğru sınıflama yaptığını belirlemek için bazı doğruluk ölçütleri bulunmaktadır. Bunlardan ikisi; test tarafından doğru tespit edilen hastaların yüzdesi olan duyarlılık ve test tarafından doğru tespit edilen sağlıklı kişilerin yüzdesi olan seçicilik kavramlarıdır.

Ayırım başarısı yüzde yüz kesin olmayan tanı testlerinin güvenilirliğinin de yüzde yüz olduğundan bahsedilemez. Bu eksiklik, tanı testi tarafından hasta veya sağlıklı olarak sınıflandırılan kişilere şüpheyle yaklaşılmasını gerektirir. Çünkü örneğin duyarlılığı %70 olan bir tanı testinin %30'luk bir hata payı bulunmaktadır. Diğer bir deyişle, testin gerçekte hasta olan bir kişiye yanlış tanı koyup sağlıklı deme ihtimali %30'dur. Duyarlılık ve seçicilik 1 değerine yani %100'e ne kadar yakınsa hata yapma olasılığı o kadar az demektir. Bu bilgilerden yola çıkılarak tanı testlerinin ayırt etme kabiliyetini saptamak adına bazı yöntemler geliştirilmiştir. Bunlardan oldukça yaygın bir biçimde kullanılan Receiver Operating Characteristic (ROC) analizidir.

Bireyde gözlenebilecek bir olayın (hastalık, ölüm, nüks vb.) erken bir zamanda belirlenmesi gerek birey için gerekse toplum için önemli bir kazanımdır. Olayın erken belirlenmesi ile bireyin yaşam kalitesinin artırılması ve bireyin sağlığı için harcanacak personel, zaman, para gibi giderlerin azaltılması sağlanabilir. Bu amaçla, birey üzerinde çeşitli zaman dilimlerinde ölçümler yapılabilir. Ölçümün olayın gerçekleşmesinden ne kadar önce yapıldığı bilgisi olayın belirlenmesinde klinisyenlere yardımcı olabilir. Bu zaman bilgisini dikkate alan ve olayın doğru belirlenmesini sağlayan analizlere başvurmak gerekir. Bu tür bir analiz ihtiyacı boylamsal verilerde ve sağkalım verilerinde ortaya çıkmaktadır. Klasik ROC analizi, nicel sonuç veren tanı testlerinin sınıflama başarısını ölçmeye yarayan bir yöntemdir. Bu analizde ölçümün yapıldığı zaman bilgisi işin içine katılmadığından değerlendirmenin doğru yapılabilmesi için yeni yöntemlere ihtiyaç duyulmuştur. Bu ihtiyaca yönelik olarak 1999 yılında Etzioni ve ark. tarafından zamana bağlı ROC analizi önerilmiştir^[1].

Literatürde, zamana bağlı veri modellerinin prediktif performansının belirlenmesi konusunun son yıllarda popülerliği artmıştır. Etzioni ve ark.'dan^[1] sonra yapılan çalışmalara bakıldığında, zamana bağlı duyarlılık, seçicilik ve ROC eğrisi kavramları Heagerty ve Zheng^[2], Zheng ve Heagerty^[3,4], Antolini ve ark.^[5] gibi araştırmacılar tarafından da ele alınmıştır. Heagerty ve ark.^[6] tarafından önerilen “kümülatif duyarlılık/dinamik seçicilik” kavramı, sabit bir t zamanında tüm bireylerin vaka ya da kontrol olarak sınıflandırılmasını temel almaktadır. Bu öneride t zamanına kadar kontrol olan bireylerin t zamanından sonra vaka olabileceği bakış açısı ile, Akritas'ın^[7] ikili dağılımın en yakın komşu kestiriminden yola çıkarak sansürlü sağkalım sürelerinin tahmini için parametrik olmayan bir yaklaşım geliştirmişlerdir.

Doğruluk ölçütlerine getirilen başka bir yeni tanım, Etzioni ve ark.^[1] ile Slate ve Turnbull^[8] tarafından önerilen “anlık (incident) duyarlılık/statik seçicilik” kavramları olmuştur. Bu yaklaşımda vakalar olayın gözlendiği zamana göre tabakalara ayrılırken (anlık), kontroller sabit bir izlem süresince olayın gözlenmediği bireyler olarak tanımlanmaktadır (statik). Bu tanımlamalar ile (olay gözlenme zamanı ortak değişken alınarak) duyarlılık ve seçiciliğin tahmin edilmesinde standart regresyon yaklaşımlarının kullanılması mümkün olmaktadır.

Heagerty ve Zheng^[2] ise çalışmalarında, bu iki farklı doğruluk ölçütü tanımlamasına “anlık duyarlılık/dinamik seçicilik” ifadesini eklemişlerdir. Yazarlara göre, bu yaklaşım ile bir birey erken zamanlarda kontrol durumunda iken, t zamanında vaka olabilmektedir. Böylece hastalık durumunun bu dinamik doğası belirtecin zamana bağlı olmasına izin vermektedir. Yazarlar, bu özellikteki doğruluk ölçütlerinin belirteç verisinin ampirik dağılımı ve hazard modeli kullanılarak kolaylıkla tahmin edilebileceğini göstermişlerdir.

Cai ve ark.^[9], sansürlü verilerde iki sınıflı testlerin parametrik modelleri üzerine çalışan Leisenring ve ark.’nın^[10] çalışmasını genişleterek, olayın gözlenme zamanının sansürlü olduğu verilerde kullanılmak üzere zamana bağlı duyarlılık ve seçicilik tahmininde yarı parametrik yöntemler ortaya koymuştur.

Zheng ve Heagerty^[4], boylamsal veriler için, yarı parametrik regresyon quantile metoduna dayanan zamana bağlı ROC eğrisi yaklaşımı öne sürmüştür. Bu metod, doğruluk ölçütlerinin zaman dışındaki ortak değişkenlere bağlı olmasına da müsaade etmektedir. Bu sayede bir belirtecin doğruluğunun hasta alt gruplarında değişiklik gösterip göstermediğini belirlemede kullanılabilir.

Rizopoulos^[11] çalışmasında Wulfsohn ve Tsiatis^[12], Henderson ve ark.^[13] ile Tsiatis ve Davidian’ın^[14] çalışmalarından yola çıkarak boylamsal verilerde bir belirtecin prediktif performansını ölçmede, biyolojik belirtecin zamana bağlı değişen doğasını da hesaba katarak verilerin ortak (joint) modellemesini önermiştir. Yazar, bu ortak modellemeyen elde edilen duyarlılık ve seçicilik kestirimlerine dayanarak hesaplanan zamana bağlı EAKA’ları, belirtecin doğruluğunun zaman içindeki değişimini belirlemede kullanmıştır. Ayrıca, tüm izlem süresince elde edilen bilgiyi özetlemek adına EAKA’ların ağırlıklı ortalamalarını hesaplayan bir ayırım indeksinin kullanılmasını önermiştir.

Pepe’nin^[15] parametrik ROC regresyon yönteminden uyarlanmış bir yarı parametrik yaklaşım geliştiren Cai ve Pepe^[16], ROC eğrisinin zamanın bir fonksiyonu olarak modellenmesi üzerine yoğunlaşarak, bu amaca yönelik bir simülasyon yöntemi önermişlerdir. Önerilen simülasyon tekniği varyans-kovaryans işlemi hakkında ayrıntılı analitik ifadelerle gerek duymamaktadır. Buna ilaveten, bootstrap gibi diğer örnekleme metodlarına göre hesaplamaları daha az zaman almaktadır. Model tanımlamadaki

koşulların azaltılması ve sağlamlığın (robust) arttırılması önerilen yaklaşımın faydalı özellikleridir.

Zamana bağlı ROC analizi ile ilgili yapılmış olan diğer çalışmalar konuya çeşitli perspektiflerden yaklaşmaktadır. Zheng ve Cai^[17], Zheng ve Heagerty^[3], Li ve Ma^[18], Hung ve Chiang^[19] ile Liu ve ark.^[20] çalışmalarında zamanın etkisi ile beraber birden fazla biyolojik belirtecin birleştirilmesini incelemiştir. Hung ve Chiang^[19], sansürlü sağkalım verisinde sansür süresinin belirteçlerle korelasyon içinde olduğu veya belirteçlerden bağımsız olduğu iki farklı durumu göz önünde bulundurarak, zamana bağlı ROC eğrileri için en uygun bileşke belirtecin belirlenmesine yönelik bir genişletilmiş GLM yöntemi önermiştir. Li ve Ma^[18] da çeşitli sansür desenleri varlığında birden fazla belirtecin birleştirilmesi üzerine zamana bağlı ROC analizi yaklaşımlarını ele almıştır. Bunların yanı sıra Li ve Ning^[21] genelde olanın aksine eksik belirteç verisi varlığında zamana bağlı ROC analizi uygulamaları ile ilgili çalışmalar yapmıştır.

Bu konu ile ilgili en yeni çalışmalardan biri 2016 yılında Martinez-Cambor ve ark.^[22] tarafından yapılmıştır. Yazarlar sağdan sansürlü veri üzerinde kümülatif/dinamik ROC eğrisi tahmini için yöntemler önermiştir. Önerilen yöntemde ilgilenilen zamandaki gerçek durumları bilinmeyen bireylere değişen gerçek durum bilgisi bir olasılık olarak atanmaktadır. Bu olasılık Cox regresyonuna dayanan yarı parametrik bir yöntem veya Kaplan-Meier yaklaşımlarına dayanan parametrik olmayan bir yöntem ile tahmin edilebilmektedir.

Yakın zamanda yapılmış bir başka çalışma ise, Karaismailoğlu^[23] tarafından yapılan doktora tezidir. Çalışmada, zamana bağlı ROC eğrisi ile tanı sürecinde kullanılan birden fazla biyobelirteç oransal hazard Cox regresyon yöntemi ile birleştirilmiştir. Elde edilen bu bileşke belirtecin sınıflama başarısı, tek tek belirteçlerle belirli bir takip süresince karşılaştırılmıştır.

Her ne kadar bu konu ile ilgili literatürde çalışmalar bulunsa da, konunun bilinirliğini arttıracak, teorik yapısını anlatan ve uygulamalarla kullanımını yaygınlaştıracak çalışmalara ihtiyaç duyulmaktadır.

Bu nedenle bu tez çalışmasının amacı;

1- Zamana bağlı ROC analizini teorik açıdan ele almak,

- 2- Yöntemin uygulanış şekline ve uygulamada karşılaşılabilecek durumlara değinmek,
- 3- Farklı alanlardan alınmış veri setleri üzerinde uygulamalar yaparak yöntemin daha iyi anlaşılmasını sağlamak,
- 4- Sağlık alanındaki kullanımı hakkında bilgiler vererek yöntemin kullanım alanının yaygınlaştırılmasında yol gösterici olmaktadır.



2. GENEL BİLGİLER

2.1. Tanı Testleri

Bir tanı testinin gerçek doğruluğu, test sonucunun kişinin gerçek durumu ile karşılaştırılmasıyla elde edilir. Bu gerçek durumun “durum mevcut” veya “durum yok” şeklinde iki gruptan oluştuğu varsayılmaktadır. Örneğin; diyabet varlığına karşılık yokluğu, iyi huylu tümöre karşılık kötü huylu tümör veya tek bir tümöre karşılık birden fazla tümör. Gerçek durum altın standart kullanılarak belirlenmektedir. Farklı testler ve uygulamalar için farklı altın standartlar kullanılmaktadır. Bunlardan yaygın olan bazıları; otopsi raporları, cerrahi operasyon bulguları, biyopsi örneklerinden elde edilen patoloji sonuçları ve mükemmel veya mükemmele yakın doğruluğa sahip diğer tanı testleridir^[24]. Altın standart veya referans test adı verilen bu testlerin hastayı sağlıklıdan ayırt etmesi yüzde yüz doğrulukla gerçekleşir. Ancak altın standart testlerin kullanımının mümkün olmadığı durumlar vardır. Örneğin; bazen altın standart testler kesin sonuç verse bile hasta açısından risk taşır veya uygulanması ekonomik değildir. Bu gibi testleri herkese uygulamak yerine önce daha az kesin testleri uygulanır. Tanı testi adı verilen bu testlerin sonucundan hareketle altın standart testi yalnızca gerçekten gerekli görülen hasta grubuna uygulamak daha doğrudur^[25]. Tanı testlerinin hastalık teşhisinde kullanılabilmesi için hasta ve sağlıklıları ne kadar doğru ayırabildikleri bilinmelidir. Bu sebeple bu amaca yönelik bazı ölçütler geliştirilmiştir.

2.2. Tanı Testlerinin Değerlendirilmesinde Kullanılan Doğruluk Ölçütleri

Bir tanı testinin doğruluğunun değerlendirilmesi için çeşitli yöntemler bulunmaktadır. Test sonucunun nitel (iki sınıflı/binary) olması durumunda kullanılan yöntemler ile sürekli ölçekte olması halinde kullanılan yöntemler farklılık gösterir. Nitel özellikteki tanı testleri için dört temel doğruluk ölçütü bulunmaktadır: duyarlılık, seçicilik, pozitif ve negatif prediktif değer. Nicel ölçümlü testlerin ise doğruluk ölçütü oldukça popüler bir yöntem olan ROC analizi ile elde edilen ROC eğrisi altında kalan alandır^[26].

2.2.1. Nitel Ölçümlü Testlerde Doğruluk

Bir tanı testinin doğruluğu çeşitli şekillerde tanımlanabilir. Bazı testler doğası gereği iki sınıflıdır (dikotom). Örneğin kişiden yalnızca evet veya hayır cevabı alınabilir. Diğer bazı testler ise ordinal veya süreklidirler ve bir eşik değeri kullanılarak pratikte dikotomize edilebilirler.

İkili değişken D hastalık durumunu göstermek için kullanılsın:

$$D = \begin{cases} 1, & \text{hasta} \\ 0, & \text{sağlıklı} \end{cases} \quad (1)$$

Ayrıca Y değişkeni de bir tanı testinin sonucunu temsil etsin. Daha büyük Y değerleri hastalığa dair daha belirleyici olsun:

$$Y = \begin{cases} 1, & \text{pozitif test sonucu} \\ 0, & \text{negatif test sonucu} \end{cases} \quad (2)$$

Çizelge 2.1. Tanı Testlerinin Performansını Değerlendirmede Kullanılan 2x2 Karar Matrisi^[25]

| | | Gerçek Durum | | Toplam |
|-------------------|--------------|--------------|--------------|--------|
| | | Hastalık Var | Hastalık Yok | |
| Tanı Testi Sonucu | Test Pozitif | a | b | a+b |
| | Test Negatif | c | d | c+d |
| Toplam | | a+c | b+d | N |

Duyarlılık bir koşullu olasılık olup, kişi gerçekte hasta iken test sonucunun pozitif olması yani kişinin test tarafından hasta olarak sınıflandırılması olasılığıdır ve $P(Y = 1 | D = 1)$ ile ifade edilir. Doğru (gerçek) pozitif oran (DPO) olarak da adlandırılır. Duyarlılık ne kadar yüksek ise testin gerçekte hasta olanları belirleme başarısı o kadar fazladır.

Testin gerçekte hasta olanları sağlıklı gruba yerleştirmesi olasılığı ise hatalı negatif oran (HNO) olarak adlandırılır ve $P(Y = 0 | D = 1)$ ile ifade edilir. Hatalı Negatif Oran aynı zamanda 1-Duyarlılık ile de hesaplanabilir.

$$\text{Doğru Pozitif Oran} = \text{Duyarlılık} = \frac{a}{a + c} \quad (3)$$

$$\text{Hatalı Negatif Oran} = 1 - \text{Duyarlılık} = \frac{c}{a + c}$$

Seçicilik de yine bir koşullu olasılık olup, kişi gerçekte sağlıklı iken test sonucunun negatif olması yani kişinin test tarafından sağlıklı olarak sınıflandırılması olasılığıdır ve $P(Y = 0 | D = 0)$ ile ifade edilir. Doğru (gerçek) negatif oran (DNO) olarak da adlandırılır. Seçicilik ne kadar yüksek ise testin gerçekte sağlıklı olanları belirleme başarısı da o kadar fazladır.

Testin gerçekte sağlıklı olanları hasta gruba yerleştirmesi olasılığı ise Hatalı Pozitif Oran (HPO) olarak adlandırılır ve $P(Y = 1 | D = 0)$ ile ifade edilir. Hatalı Pozitif Oran aynı zamanda 1-Seçicilik ile de hesaplanabilir.

$$\text{Doğru Negatif Oran} = \text{Seçicilik} = \frac{d}{b + d} \quad (4)$$

$$\text{Hatalı Pozitif Oran} = 1 - \text{Seçicilik} = \frac{b}{b + d}$$

Bahsedilen bu dört doğruluk ölçütü dışında tanı testi performansını ölçmeye yarayan diğer ölçütler Pozitif Prediktif Değer (PPD) ile Negatif Prediktif Değer (NPD) olarak tanımlanmıştır. Pozitif Prediktif Değer, tanı testi sonucu pozitif olan bireylerin gerçekte de hasta olmalarının koşullu olasılığıdır ve $P(D = 1 | Y = 1)$ ile gösterilir. Negatif Prediktif Değer ise, tanı testi sonucu negatif olan bireylerin gerçekte de sağlıklı olmalarının koşullu olasılığıdır ve $P(D = 0 | Y = 0)$ ile gösterilir.

$$\begin{aligned} \text{Pozitif Prediktif Değer} &= \frac{a}{a + b} \\ \text{Negatif Prediktif Değer} &= \frac{d}{c + d} \end{aligned} \quad (5)$$

Duyarlılık ve seçicilik ile ilgili bilinmesi gereken önemli bir özellik, hastalığın prevalansından etkilenmemeleridir. Prediktif değerler içinse aynı özellik geçerli değildir^[27]. Onlar hem hastalığın prevalansına, hem de testin performansına bağlı olarak değişkenlik gösterirler. Örneğin; düşük PPD düşük prevalansın bir sonucu olabilir veya testin hastaları yeterince iyi ayırt edememesine bağlı olabilir^[24].

2.2.2. Nicel Ölçümlü Testlerde Doğruluk

Bu kısımda ele alınan tanı testlerinin sonucu pozitif veya negatif değil, sürekli ölçekte elde edilen ölçümlerdir.

2.2.2.1. ROC Analizi

Birçok tanı testi nicel özelliğe sahiptir. Kanser tanısında kullanılan PSA veya CA-125 gibi biyolojik belirteçlerin kan serumundaki konsantrasyonları ölçülür. Karaciğer hastalıkları için ölçülen bilirubin veya böbrek yetmezliği için ölçülen kreatinin gibi diğer biyokimyasal ölçümler de de kan serumundaki konsantrasyonlarına bakılarak elde edilir.

Hücre sayımı, antikor seviyeleri, kan şekeri, kolesterol, vücut sıcaklığı ve kan basıncı tanı amacı ile yaygın olarak kullanılan klinik ölçümlerdendir^[24].

Genellikle daha yüksek D değerlerinin hastalık durumunu ifade ettiği varsayılır. Gerçekte öyle olmasa bile birtakım dönüşümler yapılarak bu varsayımın sağlanması mümkündür. Örneğin neonatal odyolojide, daha düşük signal-to-noise oranı (SNR) işitsel uyarılara verilen cevabın daha zayıf olduğunu söylemekte iken, $D = -SNR$ olarak tanımlandığında bahsedilen varsayım sağlanmış olur.

Y 'yi (test sonucu) kullanarak ikili bir karar mekanizması tanımlama medikal testlerin değerlendirilmesinde temel ilkedir; çünkü tanı testleri şu sorulara ve benzerlerine cevap arar: “Tedaviye gerek var mı, yok mu?”, “Biyopsi yapılmalı mı, yapılmamalı mı?”. Karar kuralı test sonucunun (Y) belli bir eşik değerini aşıp aşmamasına dayanır. Bu eşik değerinin belirlenmesinde etkili olan çeşitli faktörler vardır. Pepe kitabında bu faktörlere ayrıntılı olarak yer vermiştir^[24].

Sürekli ölçekte sonuç veren ve iki sınıflı hastalık tanısı koymaya yarayan tanı testlerinin hasta ile sağlıklıyı ayırt etme kabiliyetini belirlemede ROC (Receiver Operating Characteristic) analizi ve ROC eğrisi, oldukça sık kullanılan popüler bir yöntemdir.

ROC eğrisinin çizilebilmesi için, bir c eşik değeri kullanılır. Bu c eşik değeri sürekli ölçekteki test sonuçlarını iki sınıflı hale getirmeye yarar:

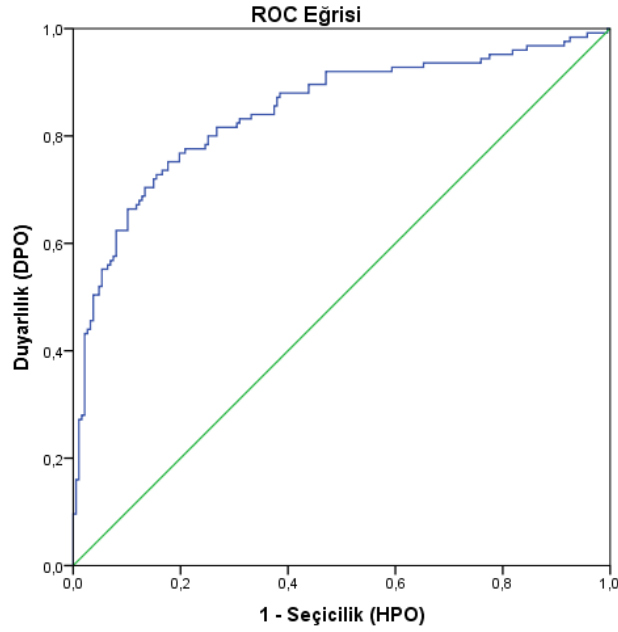
$$\begin{aligned} Y \geq c \text{ ise pozitif} \\ Y < c \text{ ise negatif} \end{aligned} \quad (6)$$

Doğru ve hatalı pozitif oranları da bu c eşik değerine göre tanımlanacak olursa;

$$\begin{aligned} DPO(c) &= P(Y \geq c \mid D = 1) \\ HPO(c) &= P(Y \geq c \mid D = 0) \end{aligned} \quad (7)$$

ROC eğrisi tüm olası c eşik değerlerinden elde edilecek olan DPO ve HPO değerlerinin birleştirilmesi ile elde edilir (Şekil 2.1).

$$ROC(.) = \{(HPO(c), DPO(c)), c \in (-\infty, \infty)\} \quad (8)$$



Şekil 2.1. ROC Eğrisi

Tıpta ROC eğrisi kullanılmasının birçok avantajı vardır. Birincisi, ROC eğrisi testin kendine özgü ayırt etme kapasitesini tek bir kesim noktasından bağımsız olarak tanımlar. İkincisi, ROC eğrisi farklı testlerin ayırt edicilik kapasitelerini karşılaştırmada da çok faydalıdır. Belirteçler çok farklı ölçeklerde olsalar bile ROC eğrisini bunları karşılaştırmada kullanmak uygundur. ROC eğrisinin diğer bir özelliği ise duyarlılık ve seçicilik değerleri kullanılarak çizildiği için, aynı bu değerler gibi hastalık prevalansından etkilenmemesidir ve bu özellik meta analizi çalışmaları için avantaj sağlar^[25,26].

ROC eğrisi, tanı testinden elde edilen her bir sonuç kesim noktası kabul edildiğinde, bu kesim noktalarına karşılık gelen her bir (duyarlılık, 1-seçicilik) noktasının

birleştirilmesi ile elde edilir. ROC eğrisi monoton artan bir fonksiyondur, (0,1) noktasından başlayıp (1,1) noktasında biter. Bu eğrinin altında kalan alan (EAKA) ise tanı testlerinin performansını değerlendirmede kullanılan bir ölçüttür ve [0,1] aralığında değer alır. Ancak pratikte bu aralığın [0.5, 1] aralığı olması beklenir. Bilgilendirici olmayan bir testin EAKA değeri 0.5'tir; yani hastayı sağlıklıdan ayırt etmesi tamamıyla rastgele gerçekleşir. Bu değere sahip ROC eğrisi şans eğrisi olarak adlandırılır ve en kötü ayırım kapasitesine sahip tanı testine aittir. EAKA değeri 0.5'ten büyük olan bir tanı testinin performansı bu değer 1'e yaklaştıkça artar. Mükemmel bir testin EAKA değerinin 1'e eşit olması beklenir^[25]. ROC eğrisi altında kalan alan parametrik ve parametrik olmayan yaklaşımlar ile hesaplanabilir.

Tanı koymada yeni göstergeler üzerine yapılan araştırmalar son yıllarda büyük artış göstermiştir ve bu durum ROC eğrisine gösterilen ilginin artmasına yol açmıştır. Araştırmacılar yapılan çalışmanın getirdiği ihtiyaca göre ROC analizinden farklı yöntemler türetmişlerdir. Bu yöntemlerden bir tanesi bu çalışmanın konusunu oluşturan zamana-bağlı ROC analizidir.

2.3. Zamana-Bağlı ROC Analizi

Tanı testleri çoğunlukla belirli bir hastalığın ortaya çıkışını belirlemek veya tahmin edebilmek için geliştirilmişlerdir. Kanseri, şikâyet ve belirtiler ışığında klinik olarak tanısı konduğunda genellikle oldukça ilerlemiş olur. Kan serumundaki PSA veya CA-125 konsantrasyonu gibi biyolojik belirteçler kanseri klinik başlangıcından önce teşhis etmeye yarar. Hastalığın klinik olarak başladığı zamandan ne kadar önce teşhis konur ise, tedavi o kadar başarılı olacaktır. Bu süreye "öncü zaman" (lead time) adı verilir^[24].

Standart ROC analizinde hasta veya sağlıklı olma durumu sabit kabul edilen özelliklerdir. Kişi çalışmaya alındığında ya hastadır ya sağlıklıdır ve ROC analizi bu tek sağlık statüsü bilgisi üzerinden gerçekleştirilir. Ancak her zaman böyle bir durum söz konusu olmayabilir. Bazı prospektif kohort çalışmalarında bireyin sağlık durumu çalışmanın seyri boyunca değişebilir. Örneğin; çalışmaya alındığında sağlıklı olan bir birey çalışma süresi içerisinde hastalığa yenik düşebilir^[6]. Ayrıca hastalığın ortaya

çıkması, biyolojik belirteç ölçümünün yapıldığı zaman ile eş zamanlı olmayabilir. Hastalığın klinik başlangıç noktasının zamana bağlı değişen doğasını göz ardı etmemek adına *zamana-bağlı* duyarlılık, seçicilik ve ROC eğrisi kavramları öne sürülmüştür^[17].

Zamana-bağlı ROC analizi kavramı iki tür veri üzerinde uygulanabilecek şekilde önerilmektedir.

- Boylamsal veriler
- Sağkalım verileri

2.3.1. Boylamsal Verilerde Zamana-Bağlı ROC Analizi

Boylamsal çalışmalarda, biyolojik belirteç ölçümleri tek bir birey için farklı zamanlarda birden fazla kez yapılır. Aynı bireye ait tekrarlı ölçümler birbirleri ile korelasyon içerisindedir; çünkü her bir ölçüm kendisinden sonraki ölçüm sonucunun görülme olasılığı üzerinde etkilidir^[28]. Bu tür bir biyolojik belirtecin ölçüldüğü zaman noktası hastalığın gerçekleşmesini tahmin etmede büyük rol oynar. Örneğin; yüksek kan basıncı ve kolesterol seviyesi miyokardiyal enfarktüsü (MI) için risk faktörüdür. Bu risk faktörleri gelecekteki bir MI vakası için tanı testi olarak kabul edilebilir. MI gözlenme zamanına yaklaşıldıkça yapılan ölçümler risk açısından daha belirleyici hale gelir^[24]. Dolayısıyla sonuçları zamana bağlı olarak değişkenlik gösteren böyle bir biyolojik belirtecin, olayın gelecekte gözlenmesi muhtemel olan ve olmayan bireyleri birbirinden ne kadar iyi ayırabildiği araştırılmak istenirse, zaman kavramı da hesaba katılarak işlem yapılması gerçeğe daha yakın sonuçlar verecektir.

Literatürde, zamana bağlı veri modellerinin sınıflama başarısının belirlenmesi ile ilgili daha önce yapılan çalışmalara bakıldığında, boylamsal verilerde ROC eğrisine zaman boyutunun eklenmesi ilk olarak Etzioni ve ark.^[1] tarafından ortaya konmuştur. Zamana bağlı duyarlılık, seçicilik ve dolayısıyla ROC eğrisi kavramları Heagerty ve Zheng^[2], Zheng ve Heagerty^[3,4], Antolini ve ark.^[5] gibi araştırmacılar tarafından da ele alınmıştır. Rizopoulos^[11], Wulfsohn ve Tsiatis^[12], Henderson ve ark.^[13] ve Tsiatis ve Davidian'ın^[14] çalışmalarından yola çıkarak boylamsal verilerde prediktif performansı ölçmek için biyolojik belirtecin zamana bağlı değişen doğasını da hesaba katarak

kullanılmak üzere boylamsal verilerin ve sağkalım verilerinin ortak (joint) modellemesini önermiştir.

Ortak modelleme yaklaşımı gerektiren çalışmalar dinamik bir yapıya sahiptir. Öyle ki; bir birey için boylamsal bilgi toplanırken, çalışmanın dinamik doğası gereği kişinin sağkalım olasılıklarının tahminlerini sürekli olarak güncellemek mümkündür. Bu özelliğin göz önünde bulundurulması, son ölçümden sonraki bir zaman diliminde olayı tecrübe etmek üzere olan bireyler ile bu zaman dilimini olaysız aşacak bireyleri tahmin etmek ile yoğunlukla ilişkilidir^[11].

i. birey için bir t zamanına kadarki $Y_i(t) = \{y_i(s); 0 \leq s \leq t\}$ boylamsal ölçümleri toplansın ve klinik olarak belirlenmiş (yani klinisyenin kişinin hayatta kalma şansını arttırabilmek için harekete geçebileceği) bir $(t, t + \Delta t]$ zaman diliminde oluşan olaylar ile ilgilenilsin. $Y_i(t)$ için uygun bir fonksiyon kullanılarak yüksek ve düşük riskli hastaları ayırmaya yarayan bir tahmin kuralı tanımlanabilir. Bu doğrultuda, $S_i(t, k, c) = \{y_i(s) \leq c_s; k \leq s \leq t\}$ başarı terimine dayanan bir c kesim noktaları vektörü düşünölsün. Başarısızlık olarak ise $F_i(t, k, c) = \mathbb{R}^{r(k,t)} \setminus \{y_i(s) \leq c_s; k \leq s \leq t\}$ ifadesi tanımlansın. Burada \mathbb{R}^n , n -boyutlu Öklid uzayını; $r(k, t)$, $[k, t]$ aralığında yapılmış ölçüm sayısını; c_s ise s zamanındaki kesim noktasını ifade etmektedir.

Yukarıda bahsedilen $S_i(t, k, c)$ ve $F_i(t, k, c)$ tanımlarını yaparken, Rizopoulos^[11], çalışmasında kullandığı HIV verisetinde biyolojik belirteç olarak kullanılan CD4 hücre sayısının düşük değerlerinin yüksek riski işaret etmesi özelliğini göz önünde bulundurmuştur. Bunun aksi durumlarda, belirtilen tanımlara gerekli düzeltmelerin yapılması gerekmektedir. Bu tanımlardan yola çıkılarak, duyarlılık ve seçicilik aşağıdaki gibi ifade edilmiştir:

$$\begin{aligned} Duy &= P(S_i(t, k, c) \mid T_i^* > t, T_i^* \in (t, t + \Delta t]; \theta) \\ Seç &= P(F_i(t, k, c) \mid T_i^* > t, T_i^* > t + \Delta t; \theta) \end{aligned} \quad (9)$$

Bu duyarlılık ve seçicilik tanımlamaları ile ilgili üç önemli nokta dikkate alınmalıdır. Birincisi; Heagerty ve Zheng'in^[2] genel tanımlamalarından yola çıkılarak

elde edilmiş *kümülatif* duyarlılık ve *dinamik* seçicilik söz konusudur. İkincisi; Zheng ve Heagerty^[3] gibi mevcut bulunan son ölçümü kullanmak yerine, son iki veya üç ölçümü kullanan daha kapsamlı bir tahmin kuralı tanımlanmaktadır. Son olarak; bu ölçümlerin her biri için farklı eşik değerleri hesaba katılmaktadır. Örneğin; sondan bir önceki ölçüm için kesim noktası c iken, son ölçüm için kesim noktası $0.5c$ olarak belirlenebilir ve bu, “ölçümdeki %50’lik bir azalma olayın güçlü bir göstergesidir” demektir^[11].

Tanı öncesi t zamanında çizilecek olan ROC eğrisi $ROC(u)$, “ u ” HPO değerindeki DPO değeridir.

$$ROC(u) = S_D[S_D^{-1}(u)] \quad (10)$$

S_D , t zamanındaki vaka gözlemlerinin dağılım fonksiyonun 1 'den çıkarılması, S_D^{-1} ise aynı işlemin kontrol gözlemleri için yapılması ile elde edilir. Veri seti N_D vaka ve $N_{\bar{D}}$ kontrolden oluşmaktadır. Vakalar için ölçüm verisi $\{Y_{ij}^D, j = 1, \dots, n_i, i = 1, \dots, N_D\}$, kontroller için ise $\{Y_{ij}^{\bar{D}}, j = 1, \dots, n_i, i = 1, \dots, N_{\bar{D}}\}$ ile gösterilsin ve i bireyi, j ise bireyin tanı öncesinde ölçümünün yapıldığı zamanı temsil etsin. $m^D(t)$ ve $s^D(t)$ Y^D 'nin, $m^{\bar{D}}(t)$ ve $s^{\bar{D}}(t)$ ise $Y^{\bar{D}}$ 'nin t 'deki ortalama ve standart sapması olsun. Bu durumda t zamanı için aşağıda verile ifadeler yazılır:

$$\begin{aligned} ROC(u) &= S[a_0(t) + a_1(t)S^{-1}(u)] \\ a_0(t) &= [m^{\bar{D}}(t) - m^D(t)/s^D(t)] \\ a_1(t) &= s^{\bar{D}}(t)/s^D(t) \end{aligned} \quad (11)$$

Bir tanı testinin zaman içerisindeki tanı koyma özelliğinin değerlendirilmesi için tanı öncesinde çeşitli zamanlarda uygulanan testin bilgisi gerekmektedir. Örneğin; kandaki biyolojik bir belirtece dayanan testler için, büyük kohort çalışmaları dahilinde elde edilmiş kan serumları bu ihtiyacı karşılayabilirler. Seri olarak depolanmış serum bankalarının yardımı ile hastalıklı durumları seçmek, onların ve kontrol gruplarının

serumlarını geriye dönük olarak tanı öncesi çeşitli zaman aralıklarında analiz etmek ve bu zamanlarda bu testler için duyarlılık ve seçicilikleri değerlendirmek prensipte mümkündür. t zamanında elde edilen tanı testi verileri vaka ve kontroller için ulaşılabilir olduğu takdirde $a_0(t)$ ve $a_1(t)$ değerleri bu zamandaki vaka ve kontrol örneklem ortalama ve varyansları kullanılarak tahmin edilebilir. Ancak, verinin düzenli ve tüm bireylerle eşzamanlı bir şekilde toplanamamasından dolayı, genellikle vaka ve kontrollerden aynı t zamanında ölçüm verisi alınması durumu gerçek dışıdır. Bu sebeple tanıdan önce (1 veya 5 yıl öncesi gibi) ilgilenilen bir zamana ait her veriye ulaşamayabileceğinden değerlendirmede zorluklar ortaya çıkabilir.

Bu tür zorlukları aşmak için Etzioni ve ark.^[1] tarafından iki yaklaşımdan bahsedilmektedir. Bunlardan birincisi biyolojik belirteç verisinin kendisini tanıdan önceki belli bir zamanın bir fonksiyonu olarak modellemek ve bu modelden yola çıkarak bu t zamanı ile ilişkili ROC eğrisini belirlemektir. Buna alternatif olan ikinci yaklaşım ise, zamanın ROC eğrisine etkisini doğrudan modellemektir.

2.3.1.1. Verinin Modellenmesi

Bu yaklaşım ilk olarak Tosteson ve Begg^[29] tarafından radyoloji okuma çalışmaları için ordinal regresyon modelleri kullanılarak tanıtılmıştır. Verideki biyolojik belirteç bilgisini tanıdan önceki herhangi bir (t) zamanının bir fonksiyonu olarak modellemek ve ROC eğrisini bu model aracılığıyla oluşturmak yaklaşımın ana fikrini oluşturmaktadır. Daha açıklayıcı bir şekilde ifade edilecek olursa, üzerinde çalışma yapılan boylamsal veride 1. haftadan 10. haftaya kadar tekrarlı ölçümler olsun. Bazı hastalara ait 10 haftalık ölçüm bilgisinin eksiksiz bir biçimde toplan(a)madığı varsayalım. Böyle bir durum söz konusu olduğunda, herhangi bir haftaya ait ölçümü eksik olan bireylerin belirteç bilgisinin, aynı haftaya ait ölçümleri tam olan bireylerin verisinden faydalanılarak modellenir. Etzioni ve ark.'na^[1] göre lineer-büyüme-eğrisi modeli altında vakalara ait verinin modellenmesinde aşağıda verilen ifadeden faydalanılır:

$$\begin{aligned}
Y_{ij}^D &= \beta_{0i}^D + \beta_{1i}^D t_{ij} + \varepsilon_{ij}^D \\
\varepsilon_{ij}^D &\sim N(0, \sigma_D^2) \\
(\beta_{0i}^D, \beta_{1i}^D) &\sim N((\beta_0^D, \beta_1^D), V^D)
\end{aligned} \tag{12}$$

Benzer ancak tanı öncesi zaman terimi bulunmayan bir model kontroller için aşağıdaki gibi tanımlanabilir:

$$\begin{aligned}
Y_{ij}^{\bar{D}} &= \beta_{0i}^{\bar{D}} + \varepsilon_{ij}^{\bar{D}} \\
\varepsilon_{ij}^{\bar{D}} &\sim N(0, \sigma_{\bar{D}}^2) \\
\beta_{0i}^{\bar{D}} &\sim N(\beta_0^{\bar{D}}, V^{\bar{D}})
\end{aligned} \tag{13}$$

Dolayısıyla a_0 ve a_1 değerleri şu şekilde elde edilebilir:

$$\begin{aligned}
a_0(t) &= -\frac{\beta_0^D + \beta_1^D t - \beta_0^{\bar{D}}}{\sqrt{\sigma_D^2 + (1-t)V^D(1/t)}} \\
a_1(t) &= \sqrt{\frac{\sigma_{\bar{D}}^2 + V^{\bar{D}}}{\sigma_D^2 + (1-t)V^D(1/t)}}
\end{aligned} \tag{14}$$

Yukarıdaki a_0 ve a_1 ifadeleri Y_{ij}^D ve $Y_{ij}^{\bar{D}}$ 'nin marjinal dağılımları olan $N(\beta_0^D + \beta_1^D t, \sigma_D^2 + (1-t)V^D(1/t))$ ve $N(\beta_0^{\bar{D}}, \sigma_{\bar{D}}^2 + V^{\bar{D}})$ 'nin değerlendirilmesi sonucu elde edilmiştir. Rastgele etki modeli parametrelerini tahmin eden bir algoritma uygulanarak elde edilen parametre kestirimleri, a_0 ve a_1 'de yerine yazılarak t zamanındaki ROC eğrisinin kestirimi elde edilebilir.

Buraya kadar kestirimler Y 'nin kendisi kullanılarak yapılmaktaydı. Aynı işlemler şimdiye kadarkinden farklı olarak Y 'nin tahmin değerlerine ait örnek ortalaması ve

varyansı kullanılarak yapıldığında, örneklem varyanslarının tahminleri yanlı olurken model doğru olduğu sürece ortalamalar m_D ve $m_{\bar{D}}$ yansız olacaktır. $\{\hat{Y}_i^{\bar{D}}, i = 1, \dots, N_{\bar{D}}\}$ şeklinde ifade edilen kontrol tahmin değerleri düşünüldüğünde, bu tahmin değerlerine ait örneklem varyansları aşağıdaki gibi yazılmaktadır:

$$\frac{\sum_i \left(\hat{Y}_i^{\bar{D}} - \frac{\sum_i \hat{Y}_i^{\bar{D}}}{N_{\bar{D}}} \right)^2}{N_{\bar{D}} - 1} = \frac{\sum_i \left(\beta_{0i}^{\bar{D}} - \sum_i \frac{\beta_{0i}^{\bar{D}}}{N_{\bar{D}}} \right)^2}{N_{\bar{D}} - 1} \quad (15)$$

2.3.1.2. ROC Eğrisinin Modellenmesi

Etzioni ve ark.^[1] tarafından verinin modellenmesine alternatif olarak önerilen bir başka yaklaşım, zamanın etkisinin doğrudan ROC eğrisinin üzerinde modellenmesidir. Pepe'nin çalışmalarında bu yaklaşıma ayrıntılı olarak değinilmektedir^[15,30]. ROC eğrisi için genel bir yazım şekli, X bireylere özgü (yaş, cinsiyet, vb.) değişkenleri temsil edecek şekilde, aşağıda verildiği gibidir:

$$ROC(u) = g \left[\sum_k \gamma_k h_k(u) + \alpha X \right] \quad (16)$$

g ve h_k fonksiyonları ROC eğrileri $(0,1) \times (0,1)$ düzleminde monoton artan olacak şekilde seçilmektedir. Φ kümülatif dağılım fonksiyonunun 1'den çıkarılmış hali olmak üzere $g = \Phi$, $h_0(u) = 1$ ve $h_1(u) = \Phi^{-1}(u)$ seçilerek elde edilen model şu şekildedir:

$$ROC(u) = \Phi[\gamma_0 + \gamma_1 \Phi^{-1}(u) + \alpha X] \quad (17)$$

Tanıdan önceki t zamanının etkisini modellemek için yukarıdaki ifade 18'deki gibi revize edilir:

$$ROC_t(u) = \Phi[\gamma_0 + \gamma_1 \Phi^{-1}(u) + \alpha t] \quad (18)$$

Her t 'deki ROC eğrisinin binormal dağılım formunda olduğu varsayılır ve farklı zamanlardaki ROC eğrileri sabit terimden (β_{0i} 'ler) kaynaklanan lineer etkiler yoluyla birbirleriyle ilişkilidir. 18'deki eşitlikte, 14'te verilen $a_0(t)$ ve $a_1(t)$ ifadeleri yerine sırasıyla $\gamma_0 + \alpha t$ ve γ_1 bulunmaktadır. Dolayısıyla vakalara ait eğimlerin (β_{1i} 'lerin) varyansının sıfır olması halinde ROC eğrisine ait model formülü, veriye ait modelden çıkarılabilir.

Doğrudan ROC eğrisinin modellenmesinin verinin modellenmesine göre birkaç tane avantajı vardır:

- 1- Parametrik olmayan bir yaklaşım olduğundan, daha az sayıda dağılımsal varsayım sağlanmasını gerektirmektedir. Bu nedenle de daha robust olduğu söylenebilir^[1].
- 2- Daha esnek bir yöntemdir çünkü iki farklı belirtecin ayırım yeteneğinin karşılaştırılmasını mümkün kılar.
- 3- Son olarak Etzioni ve ark.'na^[1] göre kullanılan yazılımlar bakımından uygulaması daha kolaydır.

2.3.2. Sağkalım Verilerinde Zamana-Bağlı ROC Analizi

Sağkalım verilerinde ilgilenilen olay genellikle ölüm veya hastalığın nüksetmesidir. Bu tür verilerde uygulanan vaka ve kontrol seçimi yaklaşımlarından bir tanesi “anlık vakalar ve uzun dönem kontroller” (incident cases and long-term controls) şeklindedir. Leisenring ve ark.^[10] ve Balasubramanian ve Lagakos^[31] zamana bağlı duyarlılık kavramını bu şekilde tanımlamışlardır. Test sonucu bir başlangıç zamanında ($t = 0$) ölçülür ve testin gelecekteki bir t zamanına kadar gerçekleşecek olayı tahmin etme kapasitesi araştırılır. Başka bir deyişle, merak edilen şey, belirli bir $[0, t]$ takip aralığında ilgilenilen olayın ortaya çıktığı ve çıkmadığı bireyleri bir tanı testinin ne ölçüde doğru tespit ettiği.

t zamanında olayın görüldüğü bir birey vaka olarak tanımlanırsa, vakaların matematiksel ifadesi ve DPO (19)'daki gibi yazılır:

$$\begin{aligned} T = t \text{ ise, } D(t) &= 1 \\ DPO(t) &= P(Y = 1 | T = t) \end{aligned} \quad (19)$$

Kontroller ise T^∞ gibi çok büyük bir zaman içerisinde olayın görülmediği bireyler olarak seçilir. HPO'ya ve kontrollere ait matematiksel ifade (20)'deki gibidir:

$$\begin{aligned} T > T^\infty \text{ ise, } D &= 0 \\ HPO &= P(Y = 1 | T > T^\infty) \end{aligned} \quad (20)$$

Kontrollerde hiçbir zaman hastalık gözlenmediğinden, HPO'nun ifadesinde zamana bağlılık söz konusu değildir.

Pepe^[24] kitabında, (19)'da verilen duyarlılık fonksiyonunu (21)'deki gibi kübik polinom şeklinde modellemiştir ve bu modeldeki parametreler marjinal regresyon yöntemleri ile tahmin edilebilir:

$$\text{logit } DPO(t) = \text{logit } P(Y = 1 | T = t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 \quad (21)$$

Zamana bağılı duyarlılığı tahmin etmek için t zamanının ortak değişken olarak alındığı bir lojistik regresyon modeli kullanılabilir. Ancak bu yöntemin birtakım kısıtlamaları bulunmaktadır. Öncelikle ilgilenilen doğruluk ölçütünün üzerindeki zaman etkisinin parametrik bir formda olması gerekmektedir. Buna ek olarak, çalışmadaki tüm vakalar için olayın gerçekleşme zamanının gözlenmiş olması gerekir. Örneğin, prospektif çalışmalarda olay zamanı bazı bireyler için sansürlü olabilir ve bu bireyler vaka veya kontrol olarak atanmadan çalışmadan çıkartılırlar. Bunun bir sonucu olarak da bu bireylerin çalışmaya sağlayacağı kısmi bilgi göz ardı edilmiş olur. Bu tür bilgi kaybını önlemek amacı ile sansürlü verilerde kullanılmak üzere Heagerty ve ark.^[6] tarafından farklı doğruluk ölçütü tanımlamaları kullanan bir başka yaklaşım önerilmiştir.

Heagerty ve ark.^[6], çalışmalarında vakaların ve kontrollerin seçilmesinde “aralık vakaları ve kontrolleri” (interval cases and controls) tanımlamasını kullanmışlardır. Anlık vakalardaki “ t zamanında” ifadesinin yerini burada “ t zamanına kadar” ifadesi almıştır. Yani eğer bireyde t zamanından önce olay gözlenmiş ise, birey vaka olarak işaretlenir. Bireyin kontrol grubuna atanabilmesi için t zamanına kadar olayın gözlenmemiş olması gerekmektedir. Vaka ve kontrol kavramlarının bu bilgilerin ışığında matematiksel olarak ifadesi (22)’deki gibidir:

$$\begin{aligned} T \leq t \text{ ise } D(t) &= 1 \\ T > t \text{ ise } D(t) &= 0 \end{aligned} \quad (22)$$

$t = 0$ başlangıç zamanında bütün bireyler kontroldür. $[0, t]$ aralığında kontrol grubunda olay gözlendikçe, bu bireyler vaka grubuna aktarılır. DPO ve HPO, $[0, t]$ aralığındaki birikimli (cumulative) olaylara dayanılarak (23)’te verildiği gibidir^[6]:

$$\begin{aligned} DPO^C(t) &= P(Y = 1 | T \leq t) \\ HPO^C(t) &= P(Y = 1 | T > t) \end{aligned} \quad (23)$$

Olay zamanının dağılımı bilindiğinde, bu birikimli DPO ve HPO değerleri insidans değerleri kullanılarak hesaplanabilir. Aynı şekilde birikimli ROC eğrisi ROC^C de bu DPO^C ve HPO^C 'den yola çıkılarak elde edilebilir. Ancak Heagerty ve ark.^[6], bu kavramları direk olarak hesaplamak için başka bir yöntem öne sürmüşlerdir. Önerdikleri yöntemin zaman etkisi için parametrik bir modele gerek duymamak ve sansürlü veriye uyarlanabilir olmak gibi avantajları bulunmaktadır. $y \in (-\infty, \infty)$ ve $t \geq 0$ iken $DPO^C(t, y)$ 'nin $HPO^C(t, y)$ 'ye karşılık çizilen ROC eğrisi,

$$DPO^C(t, y) = P(Y = 1 | T \leq t) = \frac{P(Y > y, T \leq t)}{P(T \leq t)} \quad (24)$$

$$HPO^C(t, y) = P(Y = 1 | T > t) = \frac{P(Y > y, T > t)}{P(T > t)}$$

ifadelerindeki pay ve payda bileşenlerinin tek ve iki değişkenli sağkalım tahmin edicileri kullanılarak kestirilmesi ile elde edilir.

2.3.2.1. Kaplan-Meier Tahmin Edicisi:

Duyarlılık ve seçicilik, $S(t) = P(T > t)$ sağkalım fonksiyonu ve $S(t | X > c)$ koşullu sağkalım fonksiyonu iken Bayes' teoremine göre şöyle yazılabilir:

$$P\{X > c | D(t) = 1\} = \frac{\{1 - S(t | X > c)\}P\{X > c\}}{\{1 - S(t)\}} \quad (25)$$

$$P\{X \leq c | D(t) = 0\} = \frac{S(t | X \leq c)P\{X \leq c\}}{S(t)}$$

$S(t)$ için oldukça sık kullanılan bir parametrik olmayan kestirim, Kaplan ve Meier tarafından verilmiştir^[32]. T_n, Z_i 'nin gözlemlenen değerler $\delta_i = 1$ için tekrarsız değerleri olarak tanımlansın. Kaplan-Meier tahmin edicisi şöyle tanımlanır:

$$\hat{S}_{KM}(t) = \prod_{s \in T_n, s \leq t} \left\{ 1 - \frac{\sum_j 1(Z_j = s)\delta_j}{\sum_j 1(Z_j \geq s)} \right\} \quad (26)$$

Sağkalım fonksiyonunu tahmin etmede Kaplan-Meier tahmin edicisi sansürlü gözlemler de dâhil verideki tüm bilgiyi kullanır. Bu durumda, belirteç ortak değişkeni X 'e ait ampirik dağılım fonksiyonu ile Kaplan-Meier tahmin edicisi birleştirilerek, t zamanında duyarlılık ve seçicilik için basit bir tahmin edici $\hat{F}_X(c) = \sum 1(X_i \leq c)/n$ olmak üzere aşağıdaki gibi verilebilir:

$$\hat{P}_{KM}\{X > c | D(t) = 1\} = \frac{\{1 - \hat{S}_{KM}(t | X > c)\} \{1 - \hat{F}_X(c)\}}{\{1 - \hat{S}_{KM}(t)\}} \quad (27)$$

$$\hat{P}_{KM}\{X \leq c | D(t) = 0\} = \frac{\hat{S}_{KM}(t | X \leq c) \hat{F}_X(c)}{\hat{S}_{KM}(t)}$$

Bu tahmin edici ile ilgili bir problem duyarlılık ve seçicilik monotonluğunu garantilememesidir. Tanım gereği, $c' > c$ için $P\{X > c | D(t) = 1\} \geq P\{X > c' | D(t) = 1\}$ 'dir. Ancak Bayes teoremi ve Kaplan-Meier yöntemi ile elde edilen tahmin ediciler bu monotonluğu bozabilir. Kaplan-Meier yöntemine dayalı ROC kestirimindeki ikinci problem ise koşullu Kaplan-Meier tahmin edicisi $\hat{S}_{KM}(t | X > c)$ 'nin sansür işleminin X 'ten bağımsız olduğunu varsaymasıdır.

2.3.2.2. İkili Dağılımın En Yakın Komşu Tahmin Edicisi (NNE):

Heagerty ve ark.'nın^[6] referans gösterdiği üzere, Akritas^[7] tarafından verilmiş ikili dağılım fonksiyonu $F(c, t) = P(X \leq c, T \leq t)$ veya eşleniği $S(c, t) = P(X > c, T > t)$ tahmin edicisi kullanılarak geçerli bir ROC eğrisi kestirimi elde edilebilir. Bu kestirim $F_X(s)$ 'nin X 'in dağılım fonksiyonu olduğu durumda $S(c, t) = \int_c^\infty S(t|X = s)dF_X(s)$ ifadesine bağlıdır. Akritas'ın^[7] gösterdiği gibi, kestirim şu şekildedir:

$$\hat{S}_{\lambda_n}(c, t) = \frac{1}{n} \sum_i \hat{S}_{\lambda_n}(t|X = X_i)1(X_i > c) \quad (28)$$

burada $\hat{S}_{\lambda_n}(t|X = X_i)$, λ_n düzgünleştirme parametresi ile karakterize olmuş koşullu sağkalım fonksiyonunun uygun bir tahmin edicisidir.

X 'in kategorik olmadığı ve her X_i değerinde yeterli gözlem bulunmadığı durumda $\hat{S}_{\lambda_n}(t|X = X_i)$ ifadesini tahmin etmek için birtakım düzgünleştirme işlemleri gerekecektir. Bahsedilen düzgünleştirme işlemleri için düzgünleştirme parametresi λ_n 'ye bağlı bir çekirdek (kernel) fonksiyonu olan $K_{\lambda_n}(X_j, X_i)$ 'ye başvurulur ve Kaplan-Meier kestirimi ağırlıklandırılarak aşağıdaki ifade elde edilir:

$$\hat{S}_{\lambda_n}(t|X = X_i) = \prod_{s \in T_n, s \leq t} \left\{ 1 - \frac{\sum_j K_{\lambda_n}(X_j, X_i)1(Z_j = s)\delta_i}{\sum_j K_{\lambda_n}(X_j, X_i)1(Z_j \geq s)} \right\} \quad (29)$$

Akritas^[7] tarafından çekirdek fonksiyonu olarak $(0,1)$ noktasına en yakın komşu olan $K_{\lambda_n}(X_j, X_i) = \mathbf{1}\{-\lambda_n < \hat{F}_X(X_i) - \hat{F}_X(X_j) < \lambda_n\}$ kullanılmaktadır. Burada her bir komşulukta bulunan gözlemlerin yüzdesini temsil etmek üzere $2\lambda_n \in (0,1)$ ifadesi yer almaktadır. Başka çekirdeklerin seçilmesi mümkün olsa da, en yakın komşuluğu kullanmak ROC tahminlerinin belirteç üzerindeki monoton dönüşümlerden etkilenmemesini sağlamaktadır.

Elde edilen duyarlılık ve seçicilik ise $\hat{S}_{\lambda_n}(t) = \hat{S}_{\lambda_n}(-\infty, t)$ olmak üzere şu şekildedir:

$$\hat{P}_{\lambda_n}\{X > c \mid D(t) = 1\} = \frac{1 - \hat{F}_X(c) - \hat{S}_{\lambda_n}(c, t)}{1 - \hat{S}_{\lambda_n}(t)} \quad (30)$$

$$\hat{P}_{\lambda_n}\{X \leq c \mid D(t) = 0\} = 1 - \frac{\hat{S}_{\lambda_n}(c, t)}{\hat{S}_{\lambda_n}(t)}$$

Kaplan-Meier kestiriminin aksine, yukarıdaki eşitliklerde duyarlılığın payındaki ifade c 'de monoton azalırken, seçiciliğin payı monoton artmaktadır. Bu nedenle en yakın komşu kestirimi ile ROC eğrisinin monoton artan bir yapıda olması koşulu sağlanmaktadır. NNE yaklaşımının bir başka avantajı ise, sansürün X 'e bağlı olmasına izin veriyor olmasıdır.

3. GEREÇ VE YÖNTEM

Bu çalışmada tanı testlerinin başarısının değerlendirilmesinde zamanın etkisini göz önünde bulunduran iki farklı yöntem (KM, NNE) incelenmiştir. Bu yöntemlerin dört farklı veri seti üzerinde uygulaması yapılmış ve elde edilen sonuçlar zamanı hesaba katmadan işlem yapan klasik ROC analizinin sonuçları ile karşılaştırılmıştır. Buna ek olarak, sağkalım verilerinde kullanılan bu iki zamana bağlı ROC yönteminin kendi aralarında karşılaştırılması da yapılmıştır.

Uygulamalarda Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı'nda Doç. Dr. Yaşar Sertdemir'in doktora tezinde kullanmış olduğu NSABP (National Surgical Adjuvant Breast and Bowel Project) veri tabanından alınmış kolon kanseri (c03) verisinden örneklenmiş bir alt veri seti ile Çukurova Üniversitesi Tıp Fakültesi Onkoloji Anabilim Dalı'ndan sağlanmış Hodgkin Lenfoma verisinin bir kısmı kullanılmıştır. Ayrıca R paket programında bulunan ve timeROC paketindeki uygulamalarda da kullanılan "Melano" ve "Paquid" adlı açık veri setleri üzerinde analizler yapılmıştır^[33].

Bu çalışmadaki uygulamalarda R programı versiyon 3.2.2 ve SPSS v20 kullanılmıştır^[34,35]. Zamana bağlı ROC eğrilerini çizdirmede Kaplan-Meier (KM) ve Nearest Neighbor Estimator (NNE) yöntemlerini karşılaştırmaya yarayan survivalROC ve belirlenen zaman noktaları için çizilen ROC eğrilerine ait standart hata hesaplamaları için timeROC R kütüphanelerinden faydalanılmıştır^[36,33].

3.1. Kolon Kanseri Verisi

Verinin orijinali 1024 hastadan oluşmaktadır. Uygulamalarda bu veri setinden rastgele örneklenmiş 500 hasta ile çalışılmıştır. Bu 500 hastanın 175'i çalışma sonunda ölmüş, 325 tanesi ise hayatta kalmıştır. Veride hastada tespit edilen pozitif nodül sayısı biyolojik belirteç değişkenini oluşturmaktadır. İncelenen olay ise ölümdür. Zaman değişkeni olarak ameliyattan ölüme kadarki sağkalım süresi alınmıştır.

3.2. Melano Verisi

Melano verisi 1962-1977 yılları arasında Danimarka Odense Üniversitesi Hastanesi'nde radikal operasyon geçiren 205 kötü huylu cilt kanseri hastasından elde edilen bilgilerden oluşmaktadır. Veride zaman değişkeni olarak operasyondan sonraki sağkalım süresi, ilgilenilen olay olarak ölüm, belirteç olarak ise tümör kalınlığı (1/100 mm) alınmıştır. Tüm hastalar 1977 yılının sonuna kadar takip edilmiş ve bunların 134 tanesi çalışmanın sonunda hayatta kalırken (14'ü kanserden olmamak üzere) 71 tanesi ölmüştür. Kanser nedeni olmayan bu 14 ölüm bu çalışmaya dahil edilmemiş, uygulamalar kalan 191 hastanın verisi üzerinden gerçekleştirilmiştir.

3.3. Paquid Verisi

Paquid verisi ise 1988 yılında Fransa'nın güneybatısında başlatılmış bir prospektif kohort çalışmasından elde edilen 2561 hastaya ait bilgilerden oluşmaktadır. Veride zaman değişkeni "Alzheimer başlangıcına kadar geçen süre" ve "Alzheimer başlamadan ölenlerde ölüme kadar geçen süre" olmak üzere ikiye ayrılmaktadır. Uygulamalarda ilgilenilen olay Alzheimer başlangıcı olduğundan, demans olmadan ölenler bu tez çalışmasında kullanılan veriden çıkarılmıştır. Analizler kalan 1927 birey üzerinden gerçekleştirilmiştir. Bu bireylerin 449'unda çalışma sonunda Alzheimer hastalığı ortaya çıkmıştır. Çalışmaya kayıt yaptırdıktan sonra Alzheimer başlangıcına kadar geçen süre zaman değişkenini, DSST (Digit Symbol Substitution Score Test) ve MMSE (Mini Mental State Examination) adı verilen test skorları ise iki ayrı belirteç değişkenini oluşturmaktadır. DSST ve MMSE için düşük değerler hastalığı gösterdiğinden gerekli bu değişkenlerin toplamaya göre tersi alınarak işlem yapılmıştır.

3.4. Hodgkin Lenfoma Verisi

Çukurova Üniversitesi Balcalı Hastanesi Onkoloji Anabilim Dalı'nda yapılan henüz yayınlanmamış bir çalışmadan elde edilen bu veri seti, 364 Hodgkin Lenfoma (HL) hastasından oluşmaktadır. Bu hastaların 88 tanesi çalışma sonunda HL nedeniyle ölmüş, 276'sı ise hayatta kalmıştır.

Hastalığın tanısında kullanılmak üzere geliştirilen IPS7 (International Prognostic Score); cinsiyet, yaş, hastalık evresi, hemoglobin seviyesi, akyuvar sayısı, lenfosit sayısı ve albümin seviyesi olmak üzere 7 faktörden oluşmaktadır. Ayrıca 2015 yılında yapılan bir çalışmada faktör sayısı 3'e indirgenerek IPS3 adında yeni bir skor önerilmiştir^[37]. Bu çalışmadaki uygulamalarda her iki skor için de zamana bağlı ROC analizi yapılmıştır. Zaman değişkeni olarak hastaların sağkalım süreleri alınmış, ilgilenilen olay ise ölüm olarak belirlenmiştir. IPS7 ve IPS3 skorları, biyolojik belirteç değişkenlerini oluşturmaktadır.



4. BULGULAR

4.1. Klasik ROC Yöntemi ile Zamana-Bağlı ROC Analizi Yöntemi olan Kaplan-Meier Kestirimlerinin Karşılaştırılması

4.1.1. Kolon Kanseri Verisinden Elde Edilen Bulgular

500 kolon kanseri hastasından oluşan veride, operasyon sonucunda pozitif çıkan nodül sayıları kaydedilmiştir. Verinin örnek bir görüntüsü Çizelge 4.1’de verilmiştir. Belirteç, sansür durumu ve sağkalım süresi olmak üzere üç ana değişkenden oluşmaktadır. Ölçümlerin yapıldığı zaman kesim noktalarının (3 yıldan itibaren 6 aylık ölçümler dikkate alındığı için bu noktalara ait) kümülatif frekansları Çizelge 4.2’de verilmiştir.

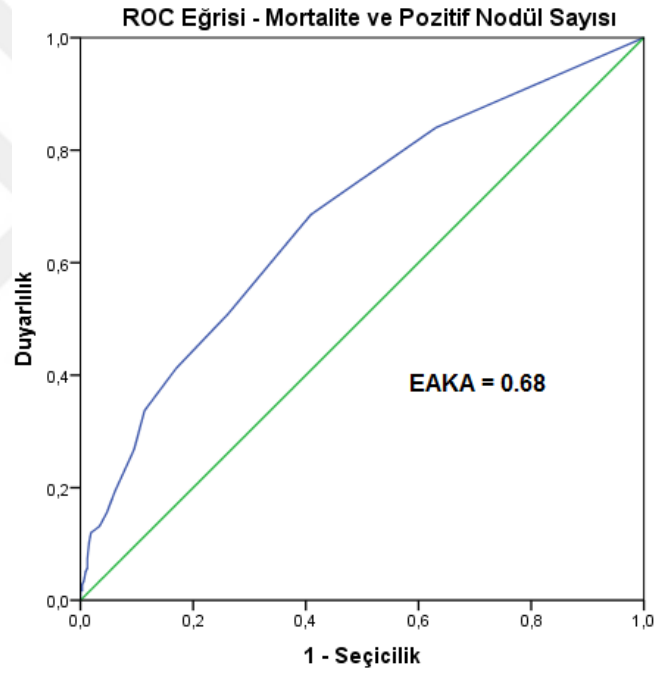
Çizelge 4.1. Kolon kanseri verisinin örnek görüntüsü

| Pozitif Nodül Sayısı | Statü | Sağkalım süresi |
|----------------------|-------|-----------------|
| 3 | 1 | 363 |
| 0 | 1 | 530 |
| 4 | 0 | 993 |
| 6 | 0 | 1093 |
| 6 | 1 | 144 |
| 11 | 1 | 493 |

Çizelge 4.2. Kolon kanseri verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları

| Zaman | Kümülatif Frekans |
|--------------|-------------------|
| 3 yıl kala | 388 |
| 2.5 yıl kala | 203 |
| 2 yıl kala | 168 |
| 1.5 yıl kala | 136 |
| 1 yıl kala | 89 |
| 6 ay kala | 38 |

Pozitif nodül sayısının tanısal değeri bulunduğundan, bu biyolojik belirtecin mortaliteyi tahmin etmedeki başarısı ile ilgilenilmektedir. Klasik ROC eğrisi çizilerek elde edilen EAKA değeri 0.68'dir (Şekil 4.1). Buna ek olarak, hastaların sağkalım süreleri, ya da başka bir deyişle "ölçümün ölümden ne kadar süre önce yapıldığı" bilgisi için içine katılarak daha iyi bir tanı performansı yakalamak amacıyla zamana bağlı ROC analizi uygulanmıştır. Hem zamana bağlı ROC yöntemi olan Kaplan-Meier kestirimi, hem de klasik ROC analizi ile elde edilen EAKA değerleri ve bunlara ait standart hatalar ile %95 güven aralıkları Çizelge 4.3'te verilmiştir.



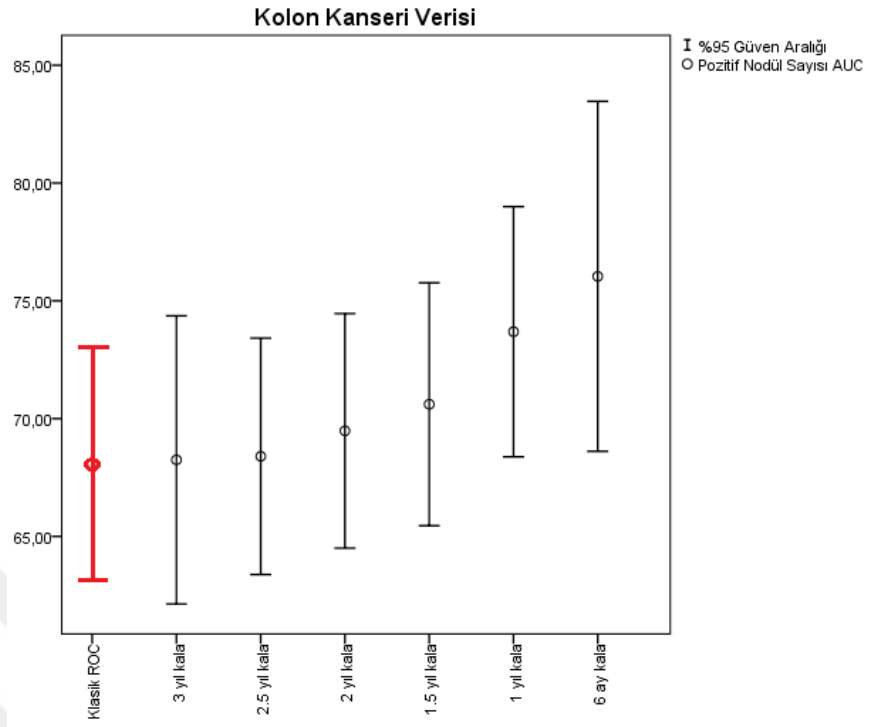
Şekil 4.1. Kolon kanseri veri seti için pozitif nodül sayısı belirtecine ait klasik ROC eğrisi

Çizelge 4.3. Kolon kanseri veri seti için 3 yıl-6 ay zaman aralığında elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|--------------|------|---------------|-------------------|
| Klasik ROC | 0.68 | 0.03 | 0.63 - 0.73 |
| 3 yıl kala | 0.68 | 0.03 | 0.62 - 0.74 |
| 2.5 yıl kala | 0.69 | 0.03 | 0.63 - 0.73 |
| 2 yıl kala | 0.69 | 0.03 | 0.65 - 0.74 |
| 1.5 yıl kala | 0.71 | 0.03 | 0.65 - 0.76 |
| 1 yıl kala | 0.74 | 0.03 | 0.68 - 0.79 |
| 6 ay kala | 0.76 | 0.04 | 0.69 - 0.83 |

Veride sağkalım sürelerinin maksimum değeri 1263 gün, yani yaklaşık üç buçuk yıldır. Bu nedenle zamana bağlı ROC analizinde ölçümün yapıldığı zaman bilgisi son 3 yıldan başlatılarak 6 aylık aralıklarla son 6 aya kadar devam ettirilmiştir. Çizelge 4.3'te görüldüğü gibi, klasik ROC eğrisine ait EAKA değeri, hastaların ölüme 3 ve 2.5 yıl kala yapılan ölçümleri ile çizdirilen ROC eğrilerinin EAKA değeri ile aynıdır. Ölüme yaklaştıkça EAKA değerleri artmış, ölüme 6 ay kala çizilen zamana bağlı ROC eğrisine ait EAKA değeri 0.76 olmuştur.

Şekil 4.2'de verilen hata çubukları grafiği incelendiğinde, 3 yıldan 1 yıla kadar zamana bağlı ROC eğrilerinin EAKA tahmin değerleri ile klasik ROC EAKA tahmin değerlerinin ortalamaları birbirlerinin güven aralıkları içine düştüğü; ancak ölüme 6 ay kala ölçülen pozitif nodül sayısının mortaliteyi tahmin yeteneğinin klasik ROC'tan daha iyi olduğu görülmüştür.



Şekil 4.2. Kolon kanseri veri seti için pozitif nodül sayısı belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği

4.1.2. Melano Verisinden Elde Edilen Bulgular

Veride 191 kötü huylu cilt kanseri hastası bulunmaktadır. Verinin örnek bir görüntüsü Çizelge 4.4'te verilmiştir. Belirteç, sansür durumu ve sağkalım süresi olmak üzere üç ana değişkenden oluşmaktadır. Ölçümlerin yapıldığı zaman kesim noktalarının (6 yıldan itibaren ölçümler dikkate alındığı için bu yıllara ait) kümülatif frekansları Çizelge 4.5'te verilmiştir.

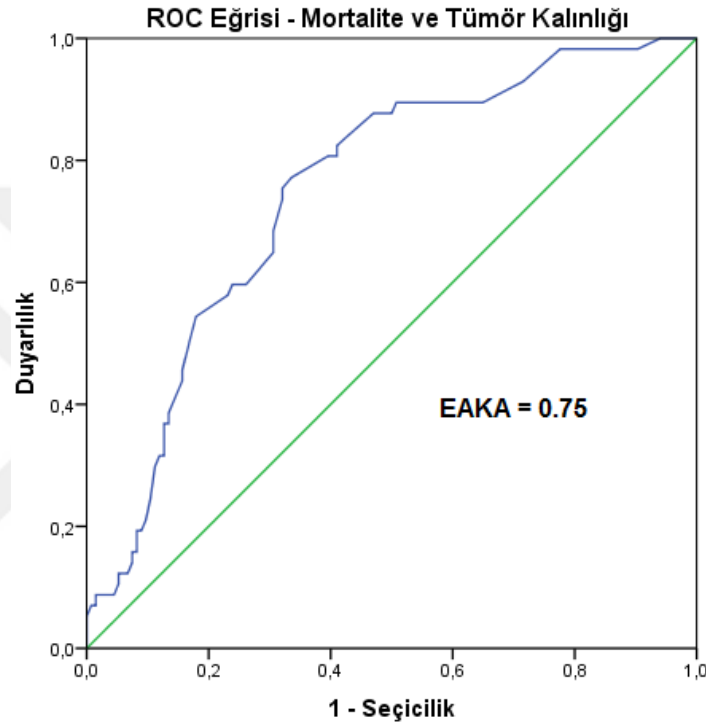
Çizelge 4.4. Melano verisinin örnek görüntüsü

| Tümör kalınlığı | Statü | Sağkalım süresi |
|-----------------|-------|-----------------|
| 12.08 | 1 | 185 |
| 4.84 | 1 | 204 |
| 1.62 | 0 | 1627 |
| 5.16 | 1 | 210 |
| 1.34 | 0 | 35 |
| 8.38 | 0 | 1508 |

Çizelge 4.5. Melano verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları

| Zaman | Kümülatif Frekans |
|------------|-------------------|
| 6 yıl kala | 111 |
| 5 yıl kala | 74 |
| 4 yıl kala | 37 |
| 3 yıl kala | 31 |
| 2 yıl kala | 16 |
| 1 yıl kala | 7 |

Hastaların operasyondan sonraki sağkalım durumları operasyon sırasında ölçülen tümör kalınlıkları ile tahmin edilmek istenmektedir. Tümör kalınlığı ölçümlerinin mortaliteyi tahmindeki başarısını belirlemek için çizilen klasik ROC eğrisi ve EAKA değeri Şekil 4.3'te verilmiştir.



Şekil 4.3. Melano veri seti için tümör kalınlığı belirtecine ait klasik ROC eğrisi

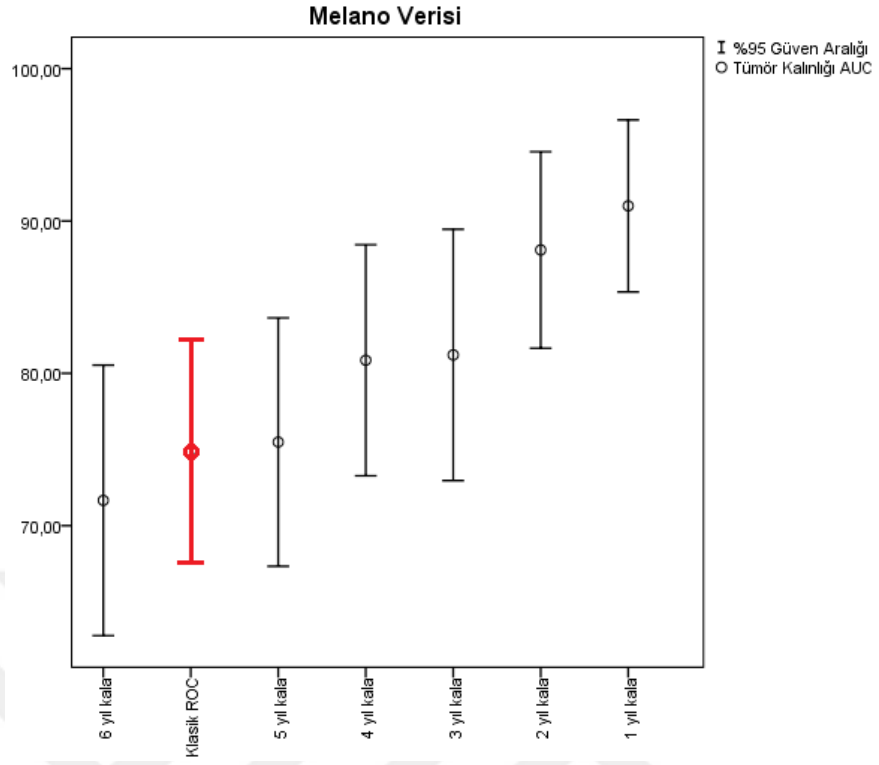
Hastalara ait sağkalım süresi maksimum 5565 gün yani yaklaşık 15 yıldır. Ancak 15 yıl, tümör kalınlığının mortaliteyi tahmin etmedeki tanısal değeri için oldukça geniş bir zaman dilimidir. Bu nedenle zamana bağlı ROC analizi uygulamasında yalnızca ölüme 6 yıl kala ve daha yakın zamandaki ölçümler dikkate alınmıştır. Son 6 yıldan başlayarak her yıla ait KM ROC analizi ile EAKA değerleri hesaplanmış, bunların ve klasik ROC eğrisinin standart hataları ve %95 güven aralıkları çizelge 4.6'da verilmiştir.

Çizelge 4.6. Melano veri seti için 6 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|------------|------|---------------|-------------------|
| 6 yıl kala | 0.73 | 0.05 | 0.63 - 0.81 |
| Klasik ROC | 0.75 | 0.04 | 0.68 - 0.82 |
| 5 yıl kala | 0.77 | 0.04 | 0.67 - 0.84 |
| 4 yıl kala | 0.80 | 0.04 | 0.73 - 0.88 |
| 3 yıl kala | 0.81 | 0.04 | 0.73 - 0.89 |
| 2 yıl kala | 0.88 | 0.03 | 0.82 - 0.95 |
| 1 yıl kala | 0.91 | 0.03 | 0.85 - 0.97 |

Çizelge 4.6'dan görüleceği üzere, ölümden 6 yıl önce yapılan ölçümler mortaliteyi tahmin etmede 0.73 başarı gösterirken, olaya yaklaşılana her yılda bu başarının arttığı gözlenmiştir. Özellikle son 2 yıl ve 1 yıl kala yapılan ölçümlerin tanı başarısı sırasıyla 0.88 ve 0.91'e çıkmıştır. Elde edilen bu bulgular klasik ROC'a ait EAKA değeri ile karşılaştırıldığında ise 1 yıldan 4 yıla kadar zamana bağlı ROC performansının klasik ROC'tan daha yüksek olduğu görülmüştür.

Şekil 4.4'te bu ROC eğrilerinin EAKA değerlerinin hata çubukları grafiği verilmiştir. Olayın gerçekleşmesine 6, 5, 4 ve 3 yıl kala yapılan ölçümlerden yola çıkılarak KM kestirimi ile elde edilen EAKA değerleri ile klasik ROC eğrisine ait EAKA değerinin güven aralıkları birbirini kapsamaktadır. Ancak dikkat edilirse 2 yıl ve 1 yıl kala değerlerine ait EAKA ortalamaları klasik ROC'un güven aralığının dışına düşmektedir. Dolayısıyla zaman bilgisinin hesaplamalara katılması ile son 2 ve son 1 yıl ölçümleri için elde edilen ROC eğrilerinin klasik ROC'a göre daha yüksek performansa sahip olduğu görülmüştür.



Şekil 4.4. Melano veri seti için tümör kalınlığı belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği

4.1.3. Paquid Verisinden Elde Edilen Bulgular

1927 hastadan oluşan veri setinin örnek görüntüsü Çizelge 4.7’de, ölçümlerin yapıldığı zaman kesim noktalarının kümülatif frekansları ise Çizelge 4.8’de verilmiştir.

Çizelge 4.7. Paquid verisinin örnek görüntüsü

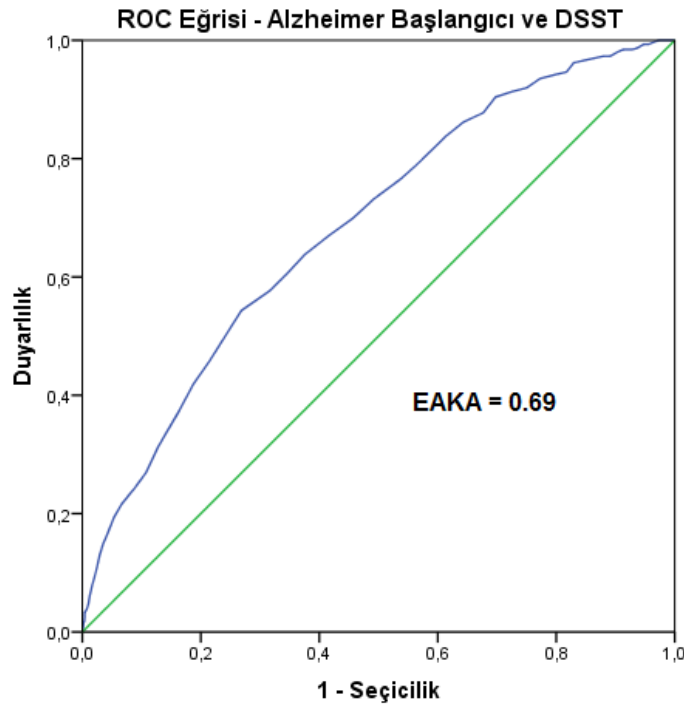
| DSST | MMSE | Statü | Sağkalım süresi |
|------|------|-------|-----------------|
| 6 | 19 | 0 | 5 |
| 22 | 25 | 1 | 6 |
| 4 | 23 | 0 | 1 |
| 15 | 28 | 1 | 2 |
| 0 | 22 | 1 | 6 |
| 12 | 22 | 1 | 11 |
| 17 | 21 | 0 | 12 |
| 11 | 19 | 1 | 4 |

Çizelge 4.8. Paquid verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları

| Zaman | Kümülatif Frekans |
|-------------|-------------------|
| 12 yıl kala | 1927 |
| 11 yıl kala | 1051 |
| 10 yıl kala | 953 |
| 9 yıl kala | 840 |
| 8 yıl kala | 734 |
| 7 yıl kala | 644 |
| 6 yıl kala | 586 |
| 5 yıl kala | 541 |
| 4 yıl kala | 370 |
| 3 yıl kala | 320 |
| 2 yıl kala | 198 |
| 1 yıl kala | 140 |

4.1.3.1. DSST Belirteç Olarak Alındığında

Bu uygulamada ilgilenilen olay hastalarda Alzheimer başlangıcı görülmesidir. Alzheimer tanısında etkili olan, dolayısıyla da biyolojik belirteç olarak kullanılan iki test skorundan bir tanesi DSST'dir (Digit Symbol Substitution Score Test). DSST'nin hastalık başlangıcını tahmindeki performansı ile ilgilenilmektedir. Zamanı için içine katmaksızın çizdirilen klasik ROC eğrisi ve bu eğriye ait EAKA değeri Şekil 4.5'te verilmiştir. "Ölçüm yapıldıktan Alzheimer başlangıcına kadar geçen süre" bilgisi kullanılarak KM yaklaşımı ile zamana bağlı ROC'lar çizdirilmiş ve EAKA değerleri Çizelge 4.9'da verilmiştir. Aynı zamanda bu değerlere ait standart hatalar ve %95 güven aralıkları aynı çizelgede sunulmuştur.



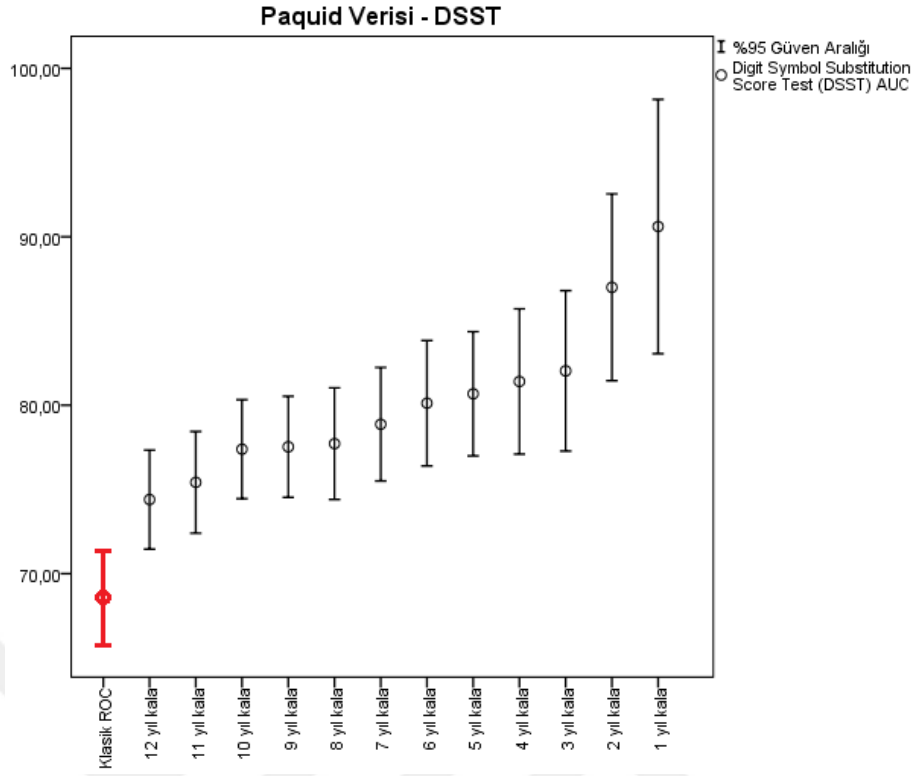
Şekil 4.5. Paquid veri seti için DSST belirteçine ait klasik ROC eğrisi

Çizelge 4.9. Paquid veri seti için DSST belirtecine 12 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|-------------|------|---------------|-------------------|
| Klasik ROC | 0.69 | 0.01 | 0.66 - 0.71 |
| 12 yıl kala | 0.74 | 0.02 | 0.71 - 0.77 |
| 11 yıl kala | 0.77 | 0.02 | 0.72 - 0.78 |
| 10 yıl kala | 0.79 | 0.02 | 0.74 - 0.80 |
| 9 yıl kala | 0.79 | 0.02 | 0.75 - 0.81 |
| 8 yıl kala | 0.79 | 0.02 | 0.74 - 0.81 |
| 7 yıl kala | 0.80 | 0.02 | 0.75 - 0.82 |
| 6 yıl kala | 0.81 | 0.02 | 0.76 - 0.84 |
| 5 yıl kala | 0.81 | 0.02 | 0.77 - 0.84 |
| 4 yıl kala | 0.83 | 0.02 | 0.77 - 0.86 |
| 3 yıl kala | 0.83 | 0.02 | 0.77 - 0.87 |
| 2 yıl kala | 0.88 | 0.03 | 0.81 - 0.93 |
| 1 yıl kala | 0.91 | 0.04 | 0.83 - 0.98 |

Hastaların maksimum sağkalım süresi 12 yıldır. Alzheimer başlangıcı için 12 yıl önce ölçümü yapılmış bir DSST skoru klinik olarak herhangi bir tanı değeri taşımasa bile, klasik ROC ile karşılaştırıldığında daha başarılı bir performans gösterdiği görülmektedir. Ölüme 12 yıl kala yapılan ölçüme ait EAKA değeri 0.74 iken, geçen her yılda bu değer artarak en son 1 yıl kala için 0.91 değerine ulaşmıştır.

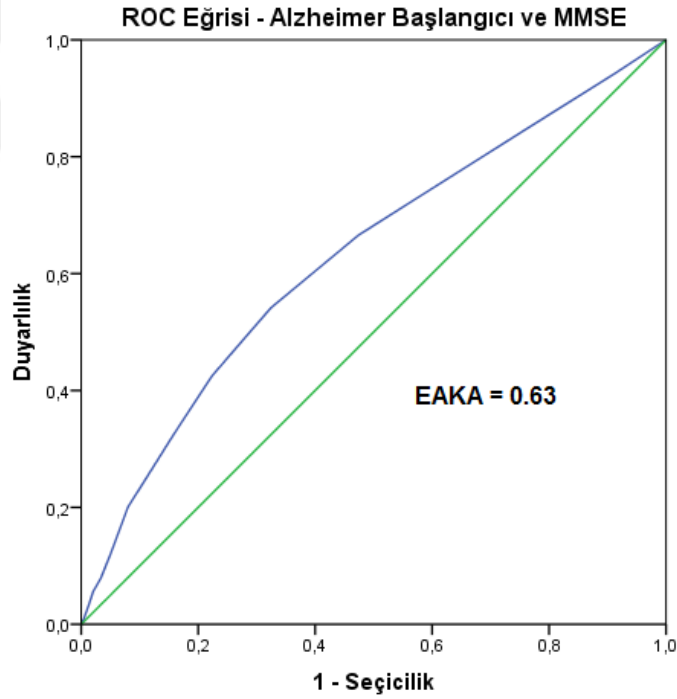
Şekil 4.6'da ölümden önceki 12 yıllık zaman dilimi boyunca her yıla ait ölçüm değerleri kullanılarak hesaplanan zamana bağlı ROC eğrisi EAKA'larına ait hata çubukları grafiği verilmiştir. Buna göre, 12 yıldan başlayıp son 1 yıla kadar olan ölçümlerin tamamı için zamana bağlı EAKA'ların, klasik ROC'un %95 güven aralığının dışında olduğu görülmüştür.



Şekil 4.6. Paquid veri seti için DSST belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği

4.1.3.2. MMSE Belirteç Olarak Alındığında

Alzheimer tanısında etkili olan, dolayısıyla da biyolojik belirteç olarak kullanılan diğer bir test skoru MMSE'dir (Mini Mental State Examination). Uygulamaların bu bölümünde, MMSE'nin hastalık başlangıcını tahmin etmedeki performansı ile ilgilenilmektedir. Zamanı işin içine katmaksızın çizdirilen klasik ROC eğrisi ve bu eğriye ait EAKA değeri Şekil 4.7'de verilmiştir. DSST belirtecinde olduğu gibi burada da "ölçüm yapıldıktan Alzheimer başlangıcına kadar geçen süre" bilgisi kullanılarak KM yaklaşımı ile zamana bağlı ROC'lar çizdirilmiş ve EAKA değerleri Çizelge 4.10'da verilmiştir. Aynı zamanda bu değerlere ait standart hatalar ve %95 güven aralıkları aynı çizelgede sunulmuştur.



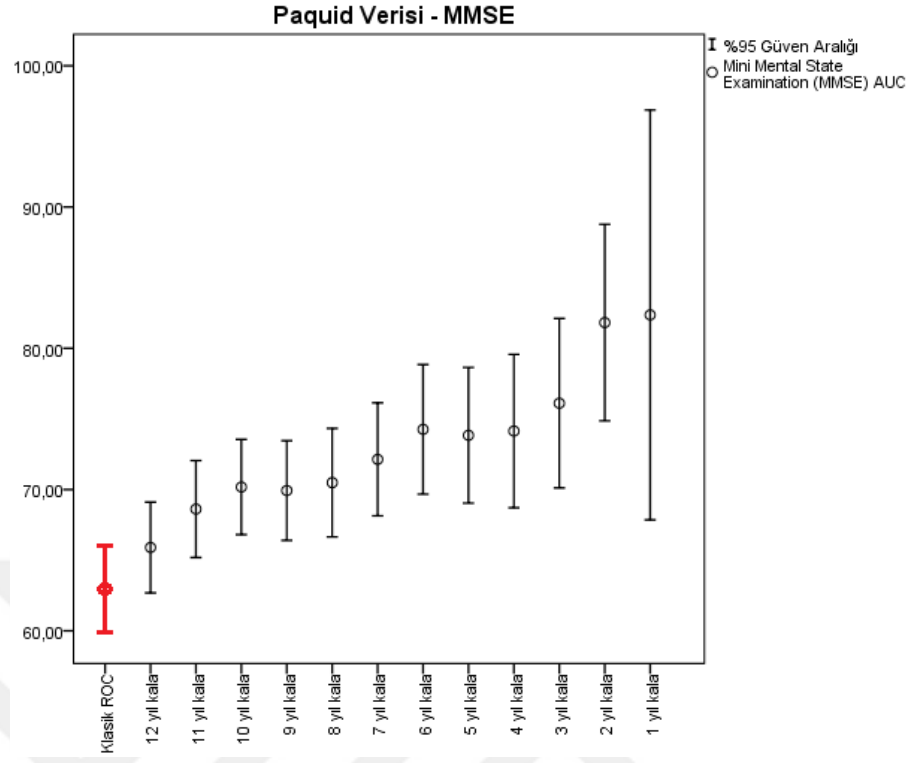
Şekil 4.7. Paquid veri seti için MMSE belirtecine ait klasik ROC eğrisi

Çizelge 4.10. Paquid veri seti için MMSE belirtecine ait 12 yılda her yıl elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|-------------|------|---------------|-------------------|
| Klasik ROC | 0.63 | 0.02 | 0.60 - 0.66 |
| 12 yıl kala | 0.66 | 0.02 | 0.63 - 0.69 |
| 11 yıl kala | 0.69 | 0.02 | 0.65 - 0.72 |
| 10 yıl kala | 0.71 | 0.02 | 0.67 - 0.74 |
| 9 yıl kala | 0.71 | 0.02 | 0.66 - 0.73 |
| 8 yıl kala | 0.71 | 0.02 | 0.67 - 0.74 |
| 7 yıl kala | 0.73 | 0.02 | 0.68 - 0.76 |
| 6 yıl kala | 0.75 | 0.02 | 0.70 - 0.79 |
| 5 yıl kala | 0.74 | 0.02 | 0.69 - 0.79 |
| 4 yıl kala | 0.74 | 0.03 | 0.69 - 0.80 |
| 3 yıl kala | 0.77 | 0.03 | 0.70 - 0.82 |
| 2 yıl kala | 0.82 | 0.04 | 0.75 - 0.89 |
| 1 yıl kala | 0.82 | 0.07 | 0.68 - 0.97 |

DSST belirtecine benzer şekilde MMSE skorunun da zamana bağlı ROC eğrilerinin EAKA'ları klasik ROC EAKA değerinden daha yüksek bulunmuştur. Ölümüne 12 yıl kala yapılan ölçüme ait EAKA değeri 0.66 iken, geçen her yılda bu değer artarak en son 1 yıl kala için 0.82 değerine ulaşmıştır.

Şekil 4.8'de ölümden önceki 12 yıllık zaman dilimi boyunca her yıla ait ölçüm değerleri kullanılarak hesaplanan zamana bağlı ROC EAKA'larına ait hata çubukları grafiği verilmiştir. Buna göre 12 yıl kala tam sınırdan başlayıp son 1 yıla kadar olan ölçümlerin tamamı için zamana bağlı EAKA'ların klasik ROC'un %95 güven aralığının dışında seyrettiği görülmüştür.



Şekil 4.8. Paquid veri seti için MMSE belirtecine ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği

4.1.4. Hodgkin Lenfoma Verisinden Elde Edilen Bulgular

1927 hastadan oluşan veri setinin örnek görüntüsü Çizelge 4.11’de, ölçümlerin yapıldığı zaman kesim noktalarının (6 yıldan itibaren 6 aylık ölçümler dikkate alındığı için bu noktalara ait) kümülatif frekansları ise Çizelge 4.12’de verilmiştir.

Çizelge 4.11. Hodgkin Lenfoma verisinin örnek görüntüsü

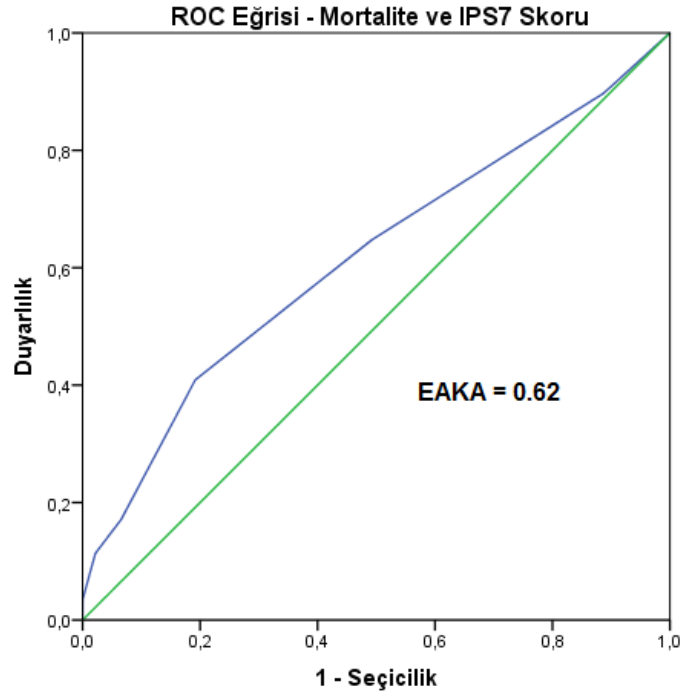
| IPS7 | IPS3 | Statü | Sağkalım süresi |
|------|------|-------|-----------------|
| 3 | 2 | 0 | 730 |
| 0 | 0 | 0 | 675 |
| 2 | 0 | 1 | 184 |
| 5 | 0 | 0 | 381 |
| 0 | 1 | 1 | 725 |
| 1 | 2 | 1 | 347 |
| 1 | 1 | 0 | 119 |
| 0 | 0 | 0 | 101 |

Çizelge 4.12. Hodgkin Lenfoma verisinde ölçümlerin yapıldığı zaman kesim noktalarındaki bireylerin kümülatif frekansları

| Zaman | Kümülatif Frekans |
|--------------|-------------------|
| 6 yıl kala | 182 |
| 5.5 yıl kala | 171 |
| 5 yıl kala | 157 |
| 4.5 yıl kala | 144 |
| 4 yıl kala | 136 |
| 3.5 yıl kala | 125 |
| 3 yıl kala | 114 |
| 2.5 yıl kala | 97 |
| 2 yıl kala | 80 |
| 1.5 yıl kala | 60 |
| 1 yıl kala | 45 |
| 6 ay kala | 18 |

4.1.4.1. IPS7 Skoru Belirteç Olarak Alındığında

364 Hodgkin Lenfoma hastasından toplanmış veri üzerinde yapılan uygulamalarda IPS7 skorunun mortaliteyi belirlemedeki başarısı araştırılmaktadır. Bu amaçla, öncelikle klasik ROC analizi ile 0.62'lik bir EAKA değeri elde edilmiş ve ROC eğrisi Şekil 4.9'da verilmiştir.



Şekil 4.9. Hodgkin Lenfoma veri seti için IPS7 skoruna ait klasik ROC eğrisi

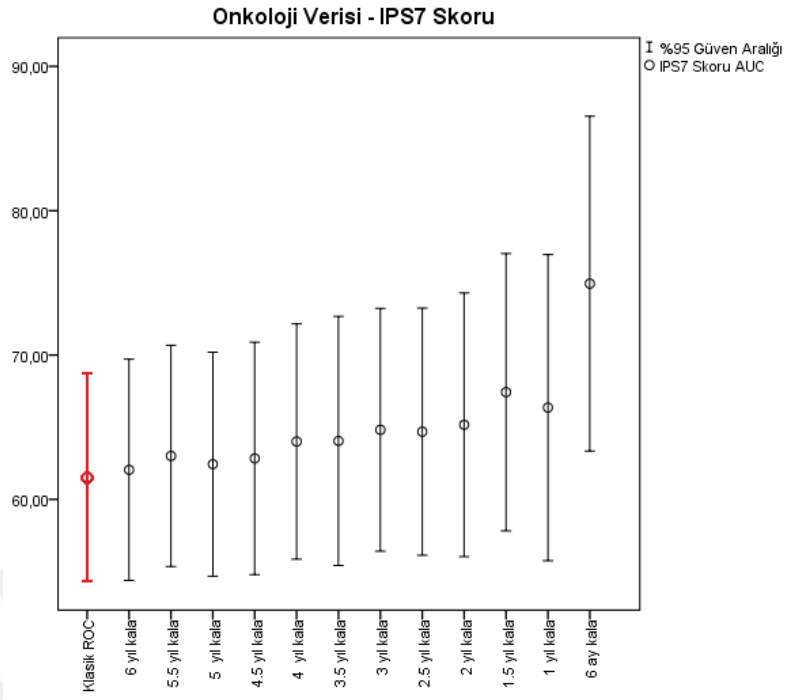
Hastaların sağkalım süreleri hesaba katılarak zamana bağlı KM kestirimi ile ROC eğrilerine ait EAKA'lar hesaplanmıştır. Bu EAKA değerleri, standart hataları ve %95 güven aralıkları ile birlikte Çizelge 4.13'te verilmiştir.

Çizelge 4.13. Hodgkin Lenfoma veri seti için IPS7 skoruna ait 6 yıl boyunca her 6 ay için elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|--------------|------|---------------|-------------------|
| Klasik ROC | 0.62 | 0.04 | 0.54 - 0.69 |
| 6 yıl kala | 0.62 | 0.04 | 0.54 - 0.70 |
| 5.5 yıl kala | 0.62 | 0.04 | 0.55 - 0.71 |
| 5 yıl kala | 0.61 | 0.04 | 0.55 - 0.70 |
| 4.5 yıl kala | 0.61 | 0.04 | 0.55 - 0.71 |
| 4 yıl kala | 0.62 | 0.04 | 0.56 - 0.72 |
| 3.5 yıl kala | 0.62 | 0.04 | 0.55 - 0.73 |
| 3 yıl kala | 0.64 | 0.04 | 0.56 - 0.73 |
| 2.5 yıl kala | 0.64 | 0.04 | 0.56 - 0.73 |
| 2 yıl kala | 0.64 | 0.05 | 0.56 - 0.74 |
| 1.5 yıl kala | 0.67 | 0.05 | 0.58 - 0.77 |
| 1 yıl kala | 0.66 | 0.05 | 0.56 - 0.77 |
| 6 ay kala | 0.75 | 0.06 | 0.63 - 0.87 |

Veride bulunan hastaların maksimum sağkalım süresi 6946 gün, yani yaklaşık 19 yıldır. IPS7 skorunun tanı koyabilmesi için 19 yıl çok geniş bir zaman aralığı olduğundan, zamana bağlı analizler yapılırken 6 yıl ve daha yakın zaman noktaları ele alınmıştır. Olayın gerçekleşmesine 6 yıl kaladan başlanarak zamana bağlı ROC analizleri her 6 ayda bir olmak üzere tekrarlanıp Çizelge 4.13'te sunulmuştur. Son 3 yıldan itibaren EAKA değerleri Klasik ROC analizi EAKA değerinden daha yüksek bulunmuştur. EAKA değerlerinde olaya yaklaşıldıkça küçük dalgalanmalar görülse de genellikle artarak son 6 ayda 0.75'e ulaşmıştır.

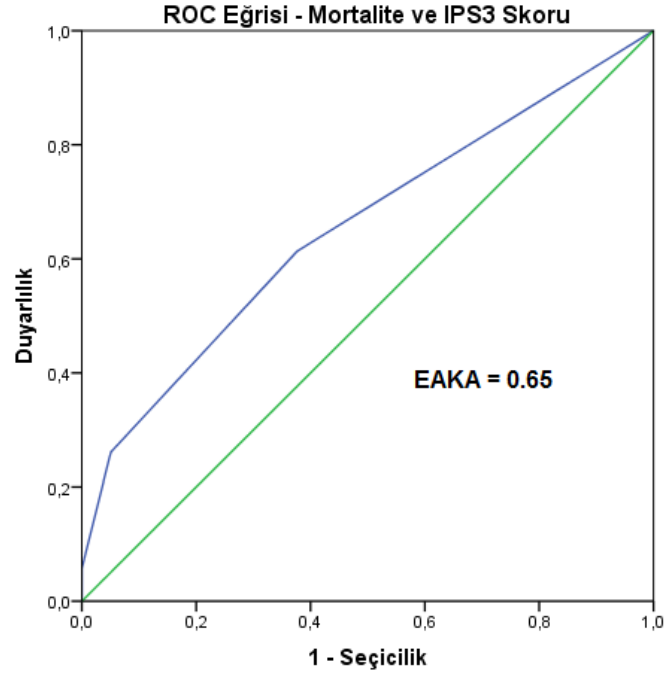
Şekil 4.10'da hata çubukları grafiği verilmiştir. Son 6 ay hariç tüm zamanlara ve klasik ROC'a ait hata çubukları birbirinin güven aralığını kapsamaktadır. Diğer bir deyişle, klasik ROC'la 6 aya kadar herhangi bir zaman için çizilen ROC eğrilerinin EAKA'ları arasında bir fark görülmemiştir. Yalnızca son 6 aydaki ölçümlerin tanı performansı daha yüksektir.



Şekil 4.10. Hodgkin Lenfoma veri seti için IPS7 skoruna ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EAKA'larının hata çubukları grafiği

4.1.4.2. IPS3 Skoru Belirteç Olarak Alındığında

364 HL hastasından toplanmış veride IPS7'ye ilave olarak IPS3 skorunun da mortaliteyi belirlemedeki başarısı araştırılmaktadır. Bu amaçla, öncelikle klasik ROC analizi ile 0.65'lik bir EAKA değeri elde edilmiş ve ROC eğrisi Şekil 4.11'de verilmiştir.



Şekil 4.11. Hodgkin Lenfoma veri seti için IPS3 skoruna ait klasik ROC eğrisi

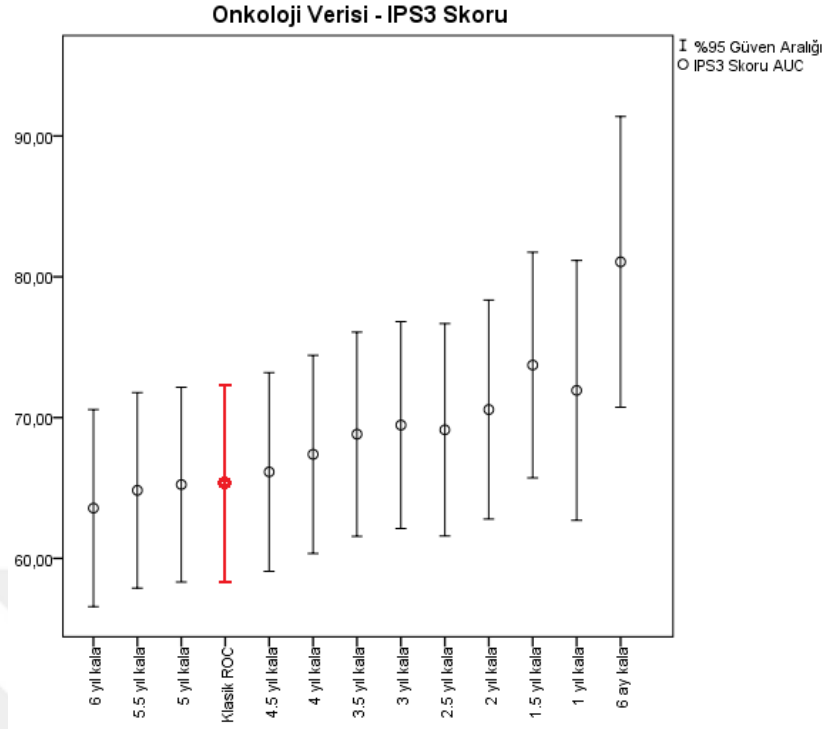
IPS7'de olduğu gibi, Hastaların sağkalım süreleri hesaba katılarak zamana bağlı KM kestirimi ile ROC eğrilerine ait EAKA'lar hesaplanmıştır. Bu EAKA değerleri, standart hataları ve %95 güven aralıkları ile birlikte Çizelge 4.14'te verilmiştir.

Çizelge 4.14. Hodgkin Lenfoma veri seti için IPS3 skoruna ait 6 yıl boyunca her 6 ay için elde edilen ROC eğrileri ile klasik ROC eğrisinin EAKA'ları, standart hata değerleri ve %95 güven aralıkları

| Zaman | EAKA | Standart Hata | %95 Güven Aralığı |
|--------------|------|---------------|-------------------|
| 6 yıl kala | 0.64 | 0.04 | 0.57 - 0.71 |
| 5.5 yıl kala | 0.65 | 0.04 | 0.58 - 0.72 |
| 5 yıl kala | 0.65 | 0.04 | 0.58 - 0.72 |
| Klasik ROC | 0.65 | 0.04 | 0.58 - 0.72 |
| 4.5 yıl kala | 0.65 | 0.04 | 0.59 - 0.73 |
| 4 yıl kala | 0.67 | 0.04 | 0.60 - 0.74 |
| 3.5 yıl kala | 0.68 | 0.04 | 0.62 - 0.76 |
| 3 yıl kala | 0.69 | 0.04 | 0.62 - 0.77 |
| 2.5 yıl kala | 0.69 | 0.04 | 0.62 - 0.77 |
| 2 yıl kala | 0.70 | 0.04 | 0.63 - 0.78 |
| 1.5 yıl kala | 0.74 | 0.04 | 0.66 - 0.82 |
| 1 yıl kala | 0.72 | 0.05 | 0.63 - 0.81 |
| 6 ay kala | 0.81 | 0.05 | 0.71 - 0.91 |

IPS7'ye benzer şekilde, burada da zamana bağlı analizler yapılırken 6 yıl ve daha yakın zaman noktaları ele alınmıştır. Olayın gerçekleşmesine 6 yıl kaladan başlanarak ölçümlere ait zamana bağlı ROC analizleri her 6 ayda bir olmak üzere tekrarlanıp Çizelge 4.14'te sunulmuştur. Klasik ROC analizi ile karşılaştırıldığında 4 yıldan öncesi ölçümler için daha iyi bir sonuç elde edildiği söylenemezken; 4. yıldan sonraki ölçümlerde zamana bağlı ROC eğrisi EAKA değerlerinin arttığı ve 6 ay kalaki ölçümler için 0.81'lik bir EAKA değerine ulaştığı görülmüştür.

Şekil 4.12'de hata çubukları grafiği verilmiştir. Son 6 ay hariç tüm zamanlara ve klasik ROC'a ait hata çubukları birbirinin güven aralığını kapsamaktadır. Yalnızca son 6 aydaki ölçümlerin tam performansının daha yüksek olduğu görülmektedir.



Şekil 4.12. Hodgkin Lenfoma veri seti için IPS3 skoruna ait farklı zamanlarda çizdirilen zamana bağlı ROC eğrileri ile klasik ROC eğrisinin EKA'larının hata çubukları grafiği

4.2. Kaplan-Meier (KM) ile En Yakın Komşu (NNE) Kestirimlerinin Karşılaştırılması

Çalışmanın bu bölümünde sağkalım verilerinde kullanılan iki zamana bağlı ROC analizi yöntemi olan Kaplan-Meier ve En Yakın Komşu kestirimlerinin, uygulamada şimdiye kadar kullanılan tüm veri setleri için farklı zamanlarda karşılaştırması yapılmıştır.

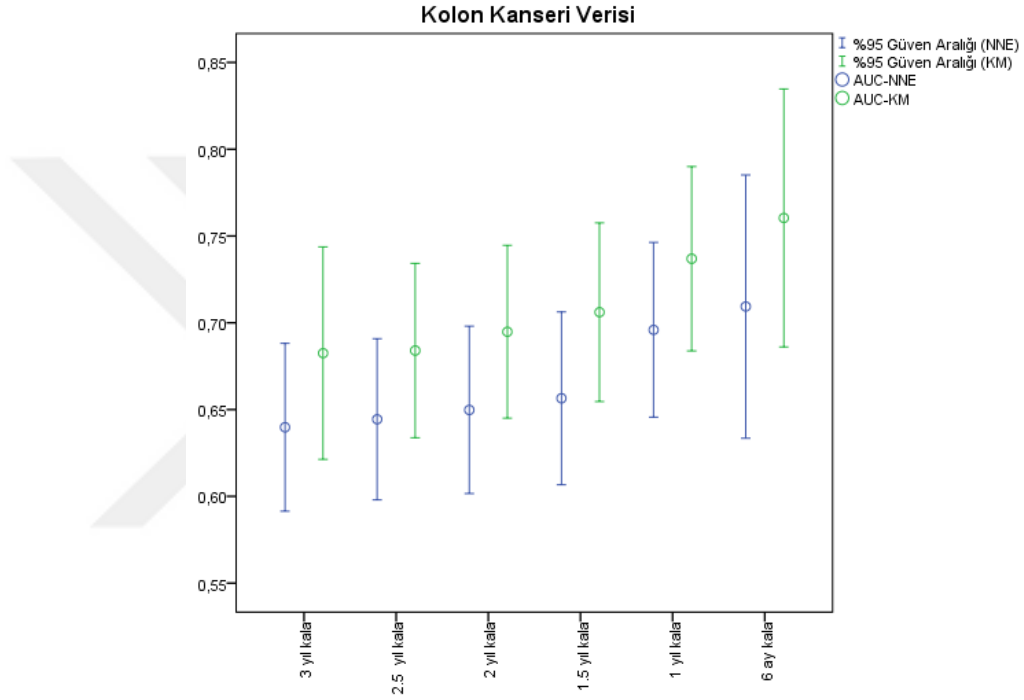
4.2.1. Kolon Kanseri Verisinden Elde Edilen Bulgular

Kolon kanseri veri seti üzerinde pozitif nodül sayısı için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.15'te verilmiştir. Zaman dilimi 3 yıldan başlayıp 6 aylık aralıklarla son 6 aya kadar devam etmektedir. Tüm zaman noktalarındaki ölçümler için KM yaklaşımına ait EAKA değerleri NNE'ninkilerden daha yüksek seyretmektedir.

Çizelge 4.15. Kolon kanseri veri seti için 3 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

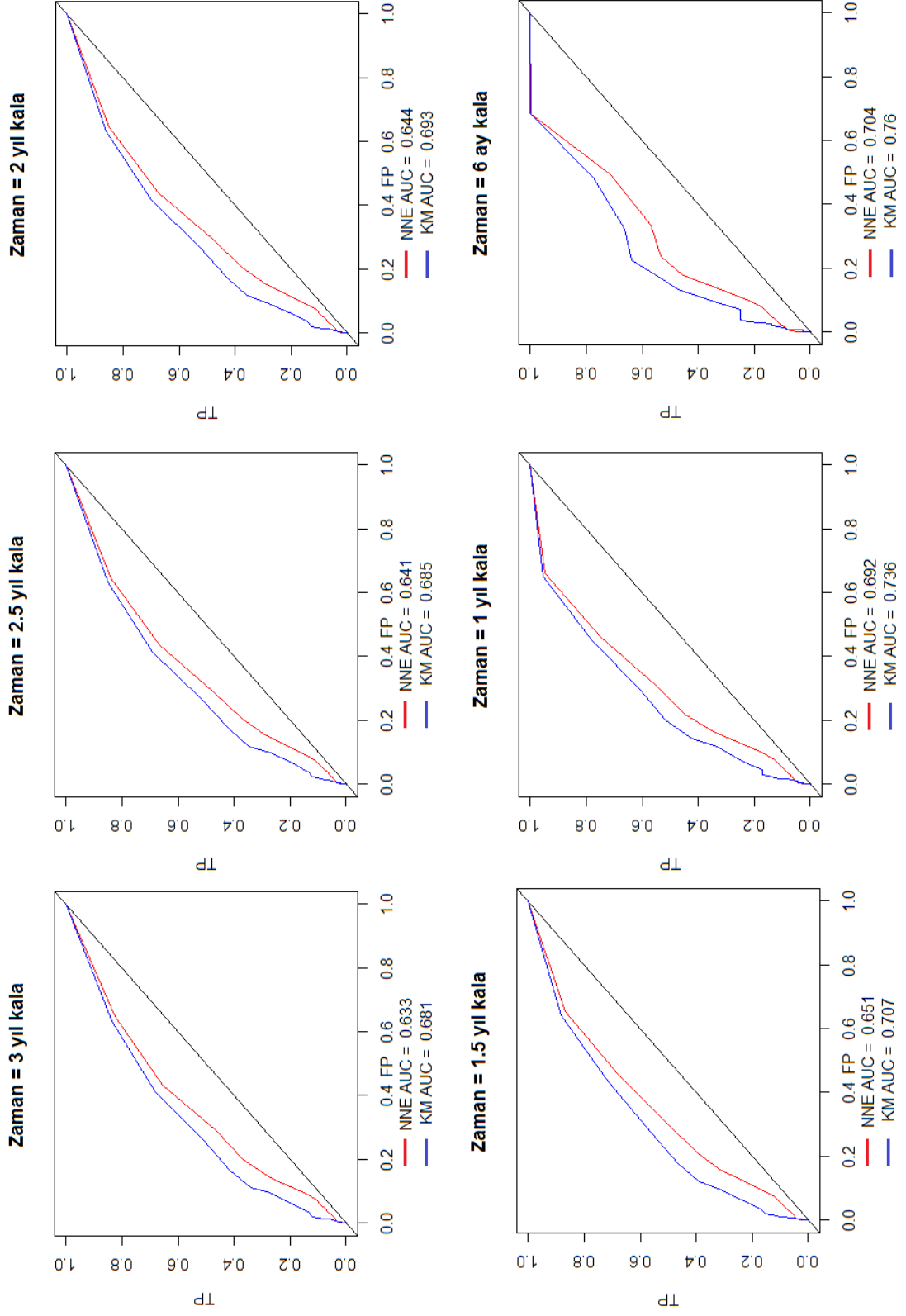
| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|--------------|---------|-------------|----------|-------------|
| 3 yıl kala | 0.68 | 0.62 - 0.74 | 0.63 | 0.59 - 0.69 |
| 2.5 yıl kala | 0.69 | 0.63 - 0.73 | 0.64 | 0.60 - 0.69 |
| 2 yıl kala | 0.69 | 0.65 - 0.74 | 0.64 | 0.60 - 0.70 |
| 1.5 yıl kala | 0.71 | 0.65 - 0.76 | 0.65 | 0.61 - 0.71 |
| 1 yıl kala | 0.74 | 0.68 - 0.79 | 0.69 | 0.65 - 0.75 |
| 6 ay kala | 0.76 | 0.69 - 0.83 | 0.70 | 0.63 - 0.79 |

Şekil 4.13'te KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. İlk dört zaman noktasında oldukça sınırda olmak üzere KM yaklaşımı NNE'den daha yüksek değerler vermiştir. Son iki zaman kesim noktasında (1 yıl ve 6 ay kala) güven aralıkları birbirini içerecek şekilde elde edilmiştir.



Şekil 4.13. Farklı zamanlardaki pozitif nodül sayısı için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği

Her iki yaklaşımdan elde edilen ROC eğrilerinin karşılaştırması Şekil 4.14'te verilmiştir.



Şekil 4.14. Kolon kanseri verisi üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin karşılaştırılması değişik zaman noktalarında karşılaştırılması

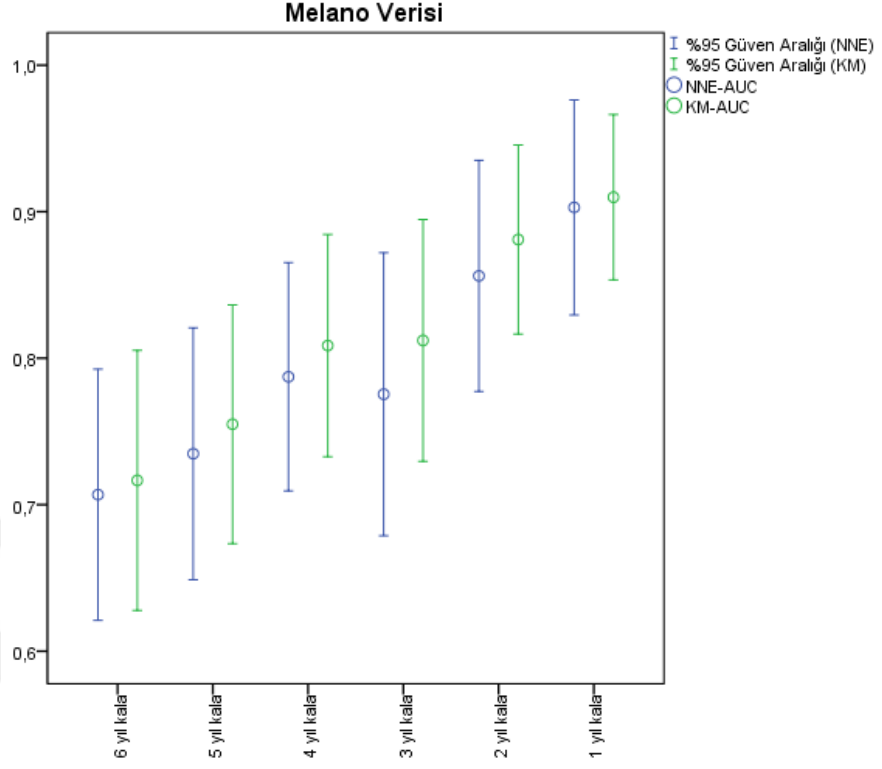
4.2.2. Melano Verisinden Elde Edilen Bulgular

Melano veri seti üzerinde tümör kalınlığı ölçümü için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.16’da verilmiştir. Zaman dilimi 6 yıldan birer yıl aralıklarla son 1 yıla kadar devam etmektedir. Tüm zaman noktalarındaki ölçümler için KM yaklaşımına ait EAKA değerleri NNE’nikilerden daha yüksek seyretmektedir.

Çizelge 4.16. Melano veri seti için 6-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

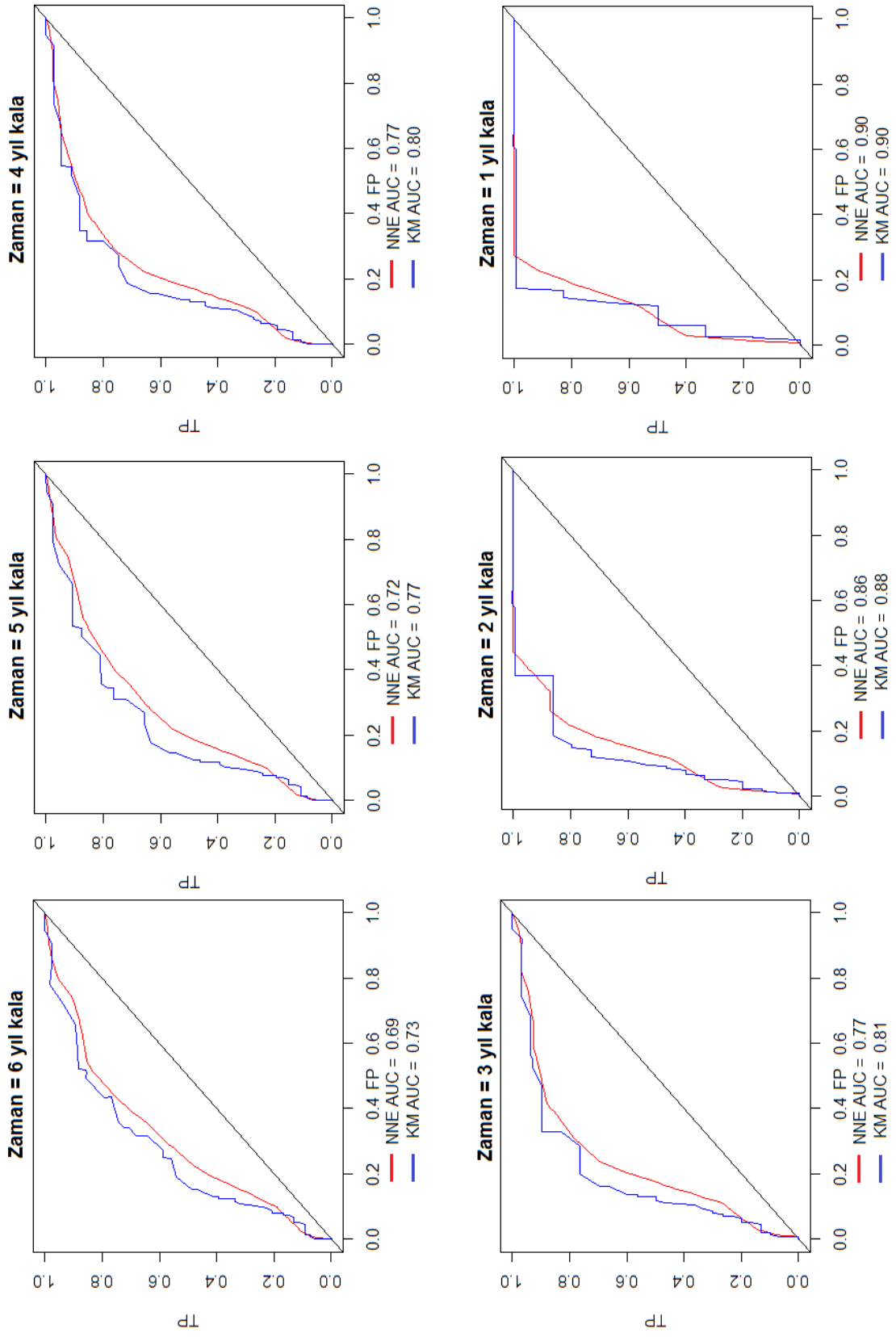
| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|------------|---------|-------------|----------|-------------|
| 6 yıl kala | 0.73 | 0.63 - 0.81 | 0.69 | 0.62 - 0.79 |
| 5 yıl kala | 0.77 | 0.67 - 0.84 | 0.72 | 0.65 - 0.82 |
| 4 yıl kala | 0.80 | 0.73 - 0.88 | 0.77 | 0.71 - 0.87 |
| 3 yıl kala | 0.81 | 0.73 - 0.89 | 0.77 | 0.68 - 0.87 |
| 2 yıl kala | 0.88 | 0.82 - 0.95 | 0.86 | 0.78 - 0.94 |
| 1 yıl kala | 0.90 | 0.85 - 0.97 | 0.90 | 0.83 - 0.98 |

Şekil 4.15’te KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. Tüm zaman kesim noktalarında yaklaşımların güven aralıklarının birbirlerini kapsadığı görülmektedir.



Şekil 4.15. Farklı zamanlardaki tümör kalınlığı ölçümleri için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği

Her iki yaklaşımdan elde edilen ROC eğrileri Şekil 4.16'da karşılaştırılmaktadır. KM yaklaşımı ile çizdirilen ROC eğrileri basamak basamak değişim gösterirken, NNE için daha düzgün eğriler elde edilmiştir. Bunun nedeni ise, NNE yaklaşımının bir düzgünleştirme parametresi içermesidir.



Şekil 4.16 Melanoma verisi üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin karşılaştırılması değişik zaman noktalarında karşılaştırılması

4.2.3. Paquid Verisinden Elde Edilen Bulgular

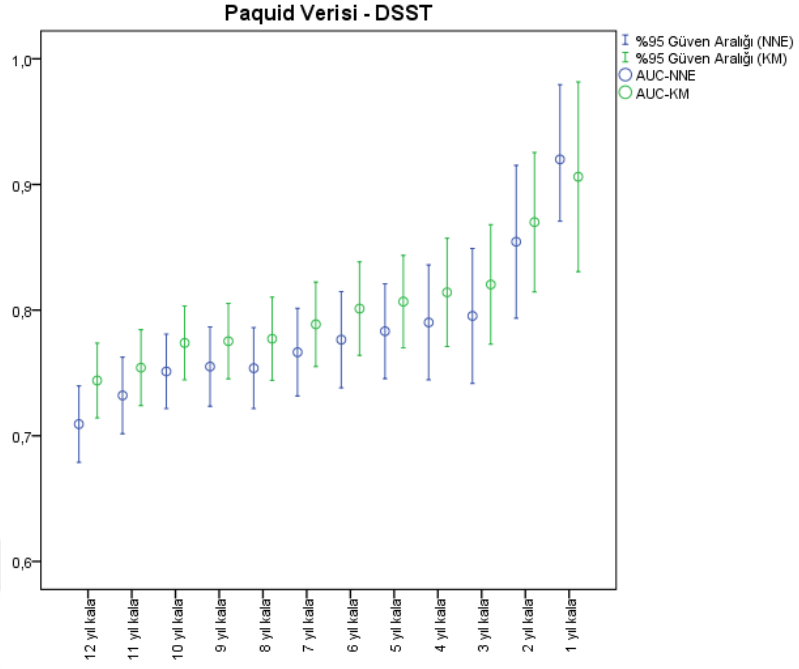
4.2.3.1. DSST Belirteç Olarak Alındığında

Paquid veri seti üzerinde DSST skorları için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.17’de verilmiştir. Zaman dilimi 12 yıldan başlayıp birer yıl aralıklarla son 1 yıla kadar devam etmektedir. Son 1 yıl kala haricinde tüm zaman noktalarındaki ölçümler için KM yaklaşımına ait EAKA değerleri NNE’nikilerden bir miktar daha yüksek seyretmektedir. Son 1 yıl kala ise yaklaşımlara ait EAKA tahminleri aynı olmuştur.

Çizelge 4.17. Paquid veri setindeki DSST skoru için 12-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

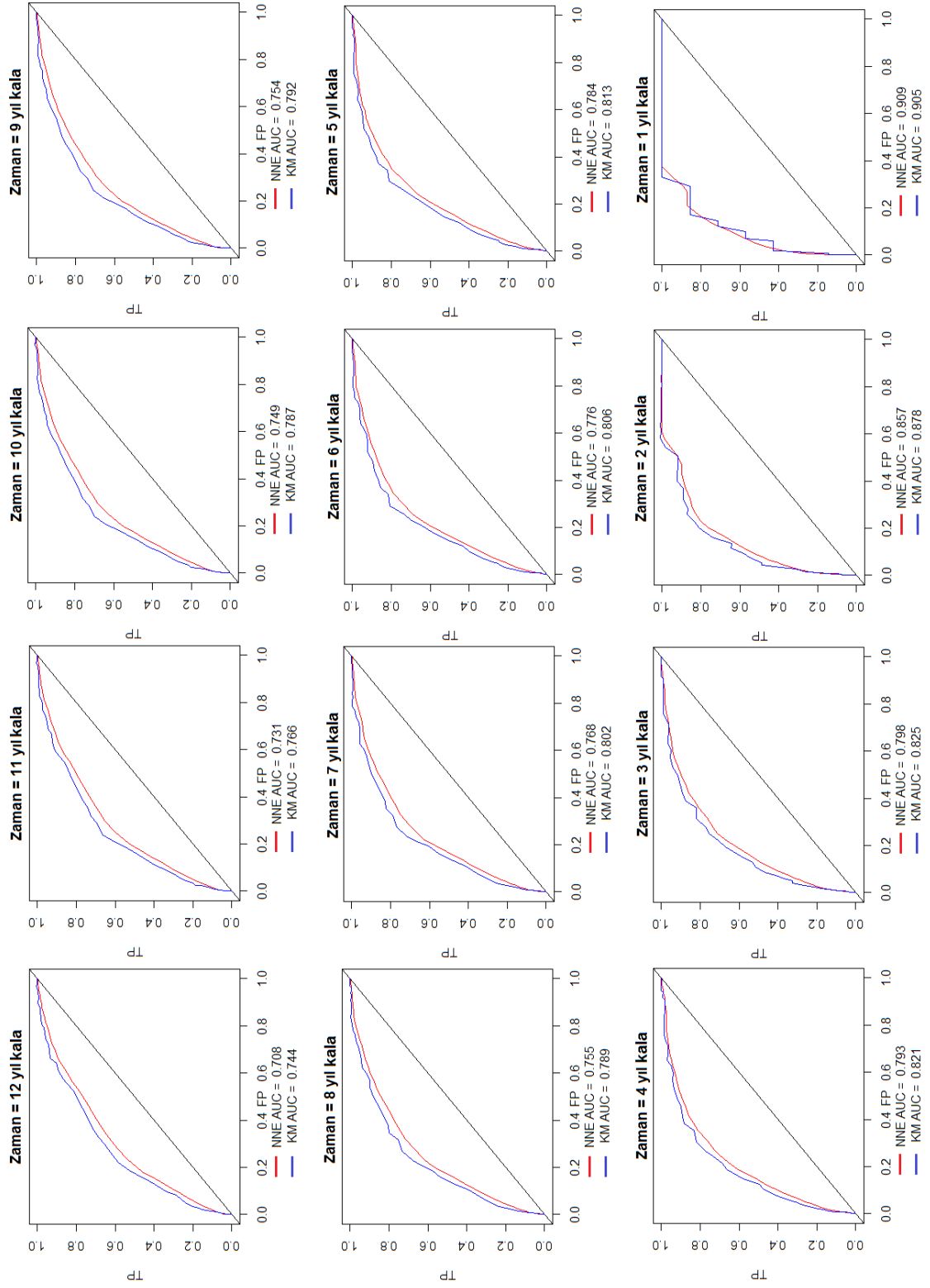
| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|-------------|---------|-------------|----------|-------------|
| 12 yıl kala | 0.74 | 0.71 - 0.77 | 0.71 | 0.71 - 0.74 |
| 11 yıl kala | 0.77 | 0.72 - 0.78 | 0.73 | 0.70 - 0.76 |
| 10 yıl kala | 0.79 | 0.74 - 0.80 | 0.75 | 0.72 - 0.78 |
| 9 yıl kala | 0.79 | 0.75 - 0.81 | 0.75 | 0.72 - 0.79 |
| 8 yıl kala | 0.79 | 0.74 - 0.81 | 0.76 | 0.72 - 0.79 |
| 7 yıl kala | 0.80 | 0.75 - 0.82 | 0.77 | 0.73 - 0.80 |
| 6 yıl kala | 0.81 | 0.76 - 0.84 | 0.78 | 0.74 - 0.81 |
| 5 yıl kala | 0.81 | 0.77 - 0.84 | 0.78 | 0.75 - 0.82 |
| 4 yıl kala | 0.83 | 0.77 - 0.86 | 0.79 | 0.74 - 0.84 |
| 3 yıl kala | 0.83 | 0.77 - 0.87 | 0.80 | 0.74 - 0.85 |
| 2 yıl kala | 0.88 | 0.81 - 0.93 | 0.86 | 0.79 - 0.92 |
| 1 yıl kala | 0.91 | 0.83 - 0.98 | 0.91 | 0.87 - 0.98 |

Şekil 4.17’de KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. 12 yıl kala değerleri dışında tüm zaman kesim noktalarında yaklaşımların güven aralıklarının birbirlerini kapsadığı görülmektedir. 12 yıl kala ölçümleri için ise KM’nin performansının NNE’den daha yüksek olduğu sonucuna varılmıştır.



Şekil 4.17. Farklı zamanlardaki DSST skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği

Her iki yaklaşımdan elde edilen ROC eğrileri Şekil 4.18'de gösterilmektedir. Önceki uygulamada da olduğu gibi burada da NNE daha düzgün ROC eğrileri vermektedir.



Şekil 4.18 Paçuid verisinde bulunan DSST belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması

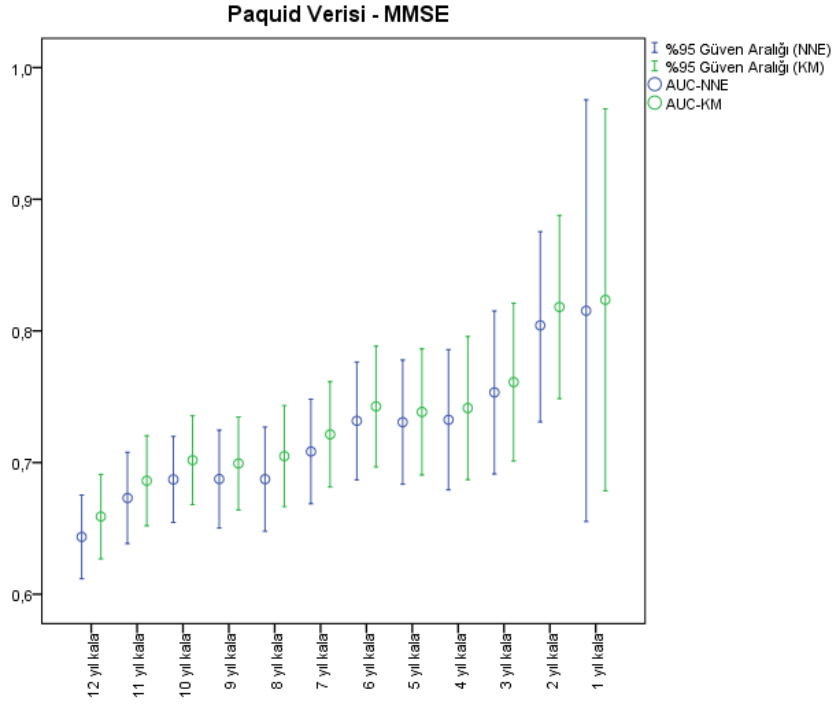
4.2.3.2. MMSE Belirteç Olarak Alındığında

Paquid veri seti üzerinde MMSE skorları için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.18’de verilmiştir. Zaman dilimi 12 yıldan başlayıp birer yıl aralıklarla son 1 yıla kadar devam etmektedir. Tüm zaman noktalarındaki ölçümler için KM yaklaşımına ait EAKA değerleri NNE’ninkilerden daha yüksek seyretmektedir.

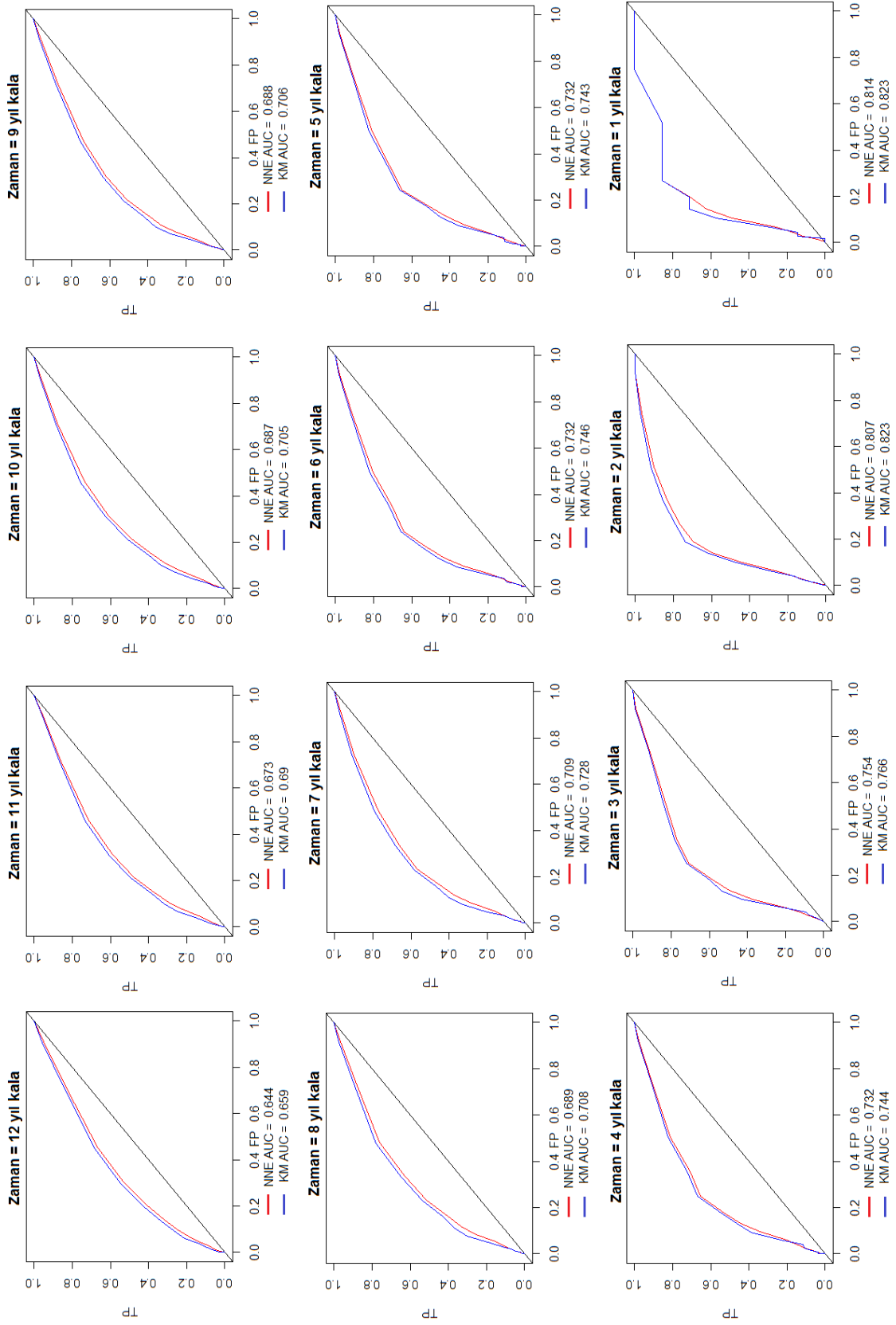
Çizelge 4.18. Paquid veri setindeki MMSE skoru için 12-1 yıl zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|-------------|---------|-------------|----------|-------------|
| 12 yıl kala | 0.66 | 0.63 - 0.69 | 0.64 | 0.61 - 0.68 |
| 11 yıl kala | 0.69 | 0.65 - 0.72 | 0.67 | 0.64 - 0.71 |
| 10 yıl kala | 0.71 | 0.67 - 0.74 | 0.69 | 0.65 - 0.72 |
| 9 yıl kala | 0.71 | 0.66 - 0.73 | 0.69 | 0.65 - 0.72 |
| 8 yıl kala | 0.71 | 0.67 - 0.74 | 0.69 | 0.65 - 0.73 |
| 7 yıl kala | 0.73 | 0.68 - 0.76 | 0.71 | 0.67 - 0.75 |
| 6 yıl kala | 0.75 | 0.70 - 0.79 | 0.73 | 0.69 - 0.78 |
| 5 yıl kala | 0.74 | 0.69 - 0.79 | 0.73 | 0.68 - 0.78 |
| 4 yıl kala | 0.74 | 0.69 - 0.80 | 0.73 | 0.68 - 0.79 |
| 3 yıl kala | 0.77 | 0.70 - 0.82 | 0.75 | 0.69 - 0.82 |
| 2 yıl kala | 0.82 | 0.75 - 0.89 | 0.81 | 0.73 - 0.88 |
| 1 yıl kala | 0.82 | 0.68 - 0.97 | 0.81 | 0.66 - 0.98 |

Şekil 4.19’da KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. Tüm zaman kesim noktalarında yaklaşımların güven aralıklarının birbirlerini kapsadığı görülmektedir. Her iki yaklaşımdan elde edilen ROC eğrileri Şekil 4.20’de gösterilmektedir.

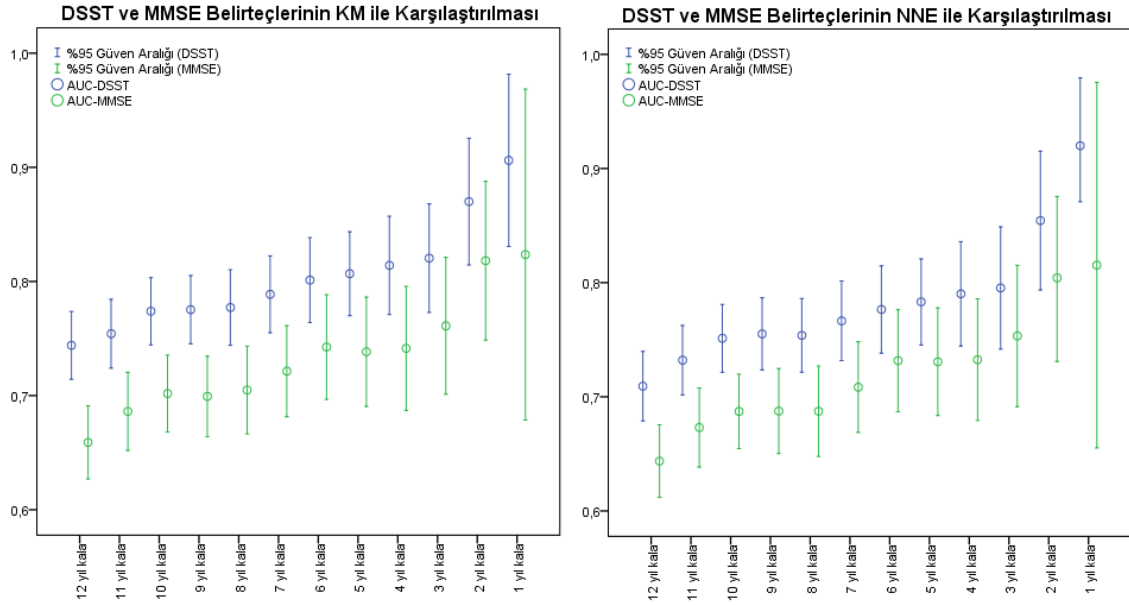


Şekil 4.19. Farklı zamanlardaki MMSE skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği



Şekil 4.20. Paçuid verisinde bulunan MMSE belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması

4.2.3.3. DSST ve MMSE Belirteçlerinin Karşılaştırılması



Şekil 4.21. Paquid veri seti için DSST ve MMSE belirteçlerinin KM ve NNE yöntemleri ile farklı zamanlarda karşılaştırılması

Paquid verisinde belirteç olarak kullanılan DSST ve MMSE skorlarının KM ve NNE yöntemleri ile tek tek karşılaştırılmasının yanı sıra, iki belirtecin birbirlerine göre performansları da karşılaştırılmıştır (Şekil 4.21). İki yöntemde de olaya en yakın birkaç yıl hariç tüm zaman noktalarında DSST belirtecinin sınıflama başarısı MMSE'den daha yüksek bulunmuştur. KM yönteminde yalnızca son 2 yıl ölçümleri için iki belirteç arasında farklılık görülmemektedir. Diğer taraftan NNE yönteminde de son 3 yıl için aynı durum söz konusudur. Buna rağmen bu zamanlarda MMSE'nin özellikle son 1 yıl ölçümlerinden elde edilen EAKA tahmininin güven aralığının DSST'den daha geniş olması DSST belirtecini daha tercih edilebilir kılabilir.

4.2.4. Hodgkin Lenfoma Verisinden Elde Edilen Bulgular

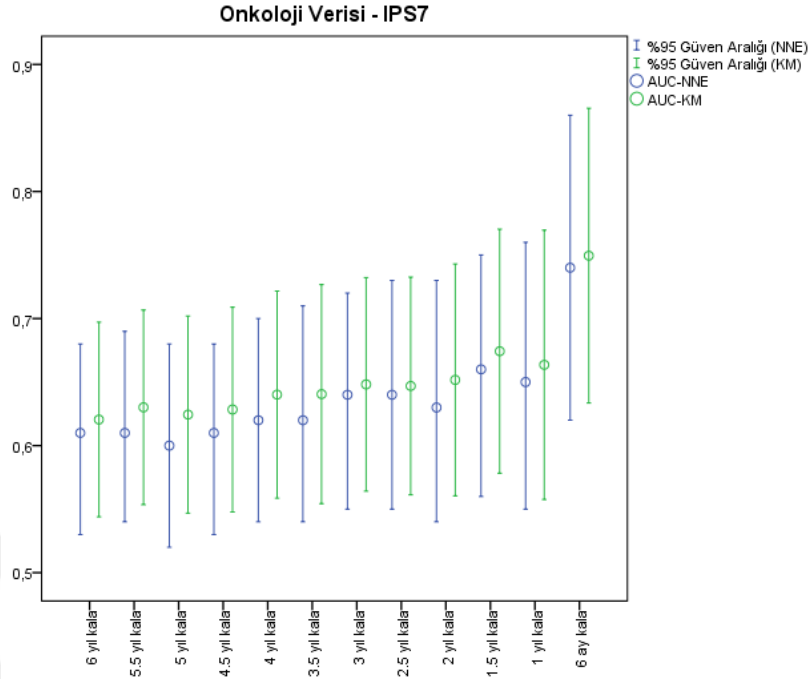
4.2.4.1. IPS7 Skoru Belirteç Olarak Alındığında

Hodgkin Lenfoma veri seti üzerinde IPS7 skorları için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.19’da verilmiştir. Zaman dilimi 6 yıldan başlayıp 6 ay aralıklarla son 6 aya kadar devam etmektedir. İki yaklaşım için tüm yıllara ait ölçümler ile elde edilen EAKA değerleri birbirine çok yakındır.

Çizelge 4.19. Hodgkin Lenfoma veri setindeki IPS7 skoru için 6 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

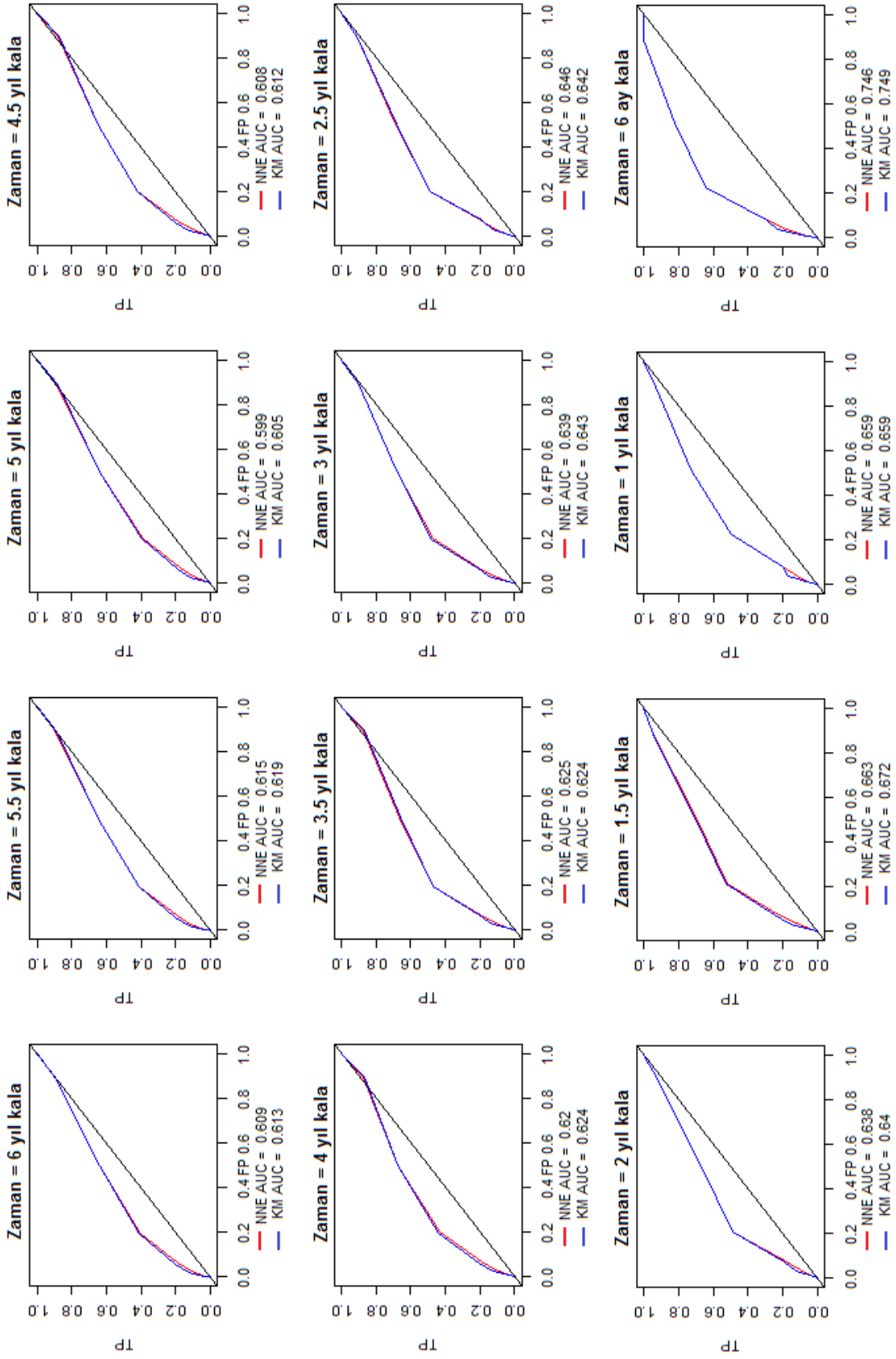
| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|--------------|---------|-------------|----------|-------------|
| 6 yıl kala | 0.61 | 0.54 - 0.70 | 0.61 | 0.53 - 0.68 |
| 5.5 yıl kala | 0.62 | 0.55 - 0.71 | 0.62 | 0.54 - 0.69 |
| 5 yıl kala | 0.61 | 0.55 - 0.70 | 0.60 | 0.52 - 0.68 |
| 4.5 yıl kala | 0.61 | 0.55 - 0.71 | 0.61 | 0.53 - 0.68 |
| 4 yıl kala | 0.62 | 0.56 - 0.72 | 0.62 | 0.54 - 0.70 |
| 3.5 yıl kala | 0.62 | 0.55 - 0.73 | 0.63 | 0.54 - 0.71 |
| 3 yıl kala | 0.64 | 0.56 - 0.73 | 0.64 | 0.55 - 0.72 |
| 2.5 yıl kala | 0.64 | 0.56 - 0.73 | 0.65 | 0.55 - 0.73 |
| 2 yıl kala | 0.64 | 0.56 - 0.74 | 0.64 | 0.54 - 0.73 |
| 1.5 yıl kala | 0.67 | 0.58 - 0.77 | 0.66 | 0.56 - 0.75 |
| 1 yıl kala | 0.66 | 0.56 - 0.77 | 0.66 | 0.55 - 0.76 |
| 6 ay kala | 0.75 | 0.63 - 0.87 | 0.75 | 0.62 - 0.86 |

Şekil 4.22’de KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. Tüm zaman kesim noktalarında yaklaşımların güven aralıklarının birbirlerini içerdiği görülmektedir.



Şekil 4.22. Farklı zamanlardaki IPS7 skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği

KM ve NNE yaklaşımlarından elde edilen ROC eğrileri karşılaştırmaları Şekil 4.23'te görülmektedir.



Şekil 4.23. Hodgkin Lenfoma verisinde bulunan IPS7 belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması

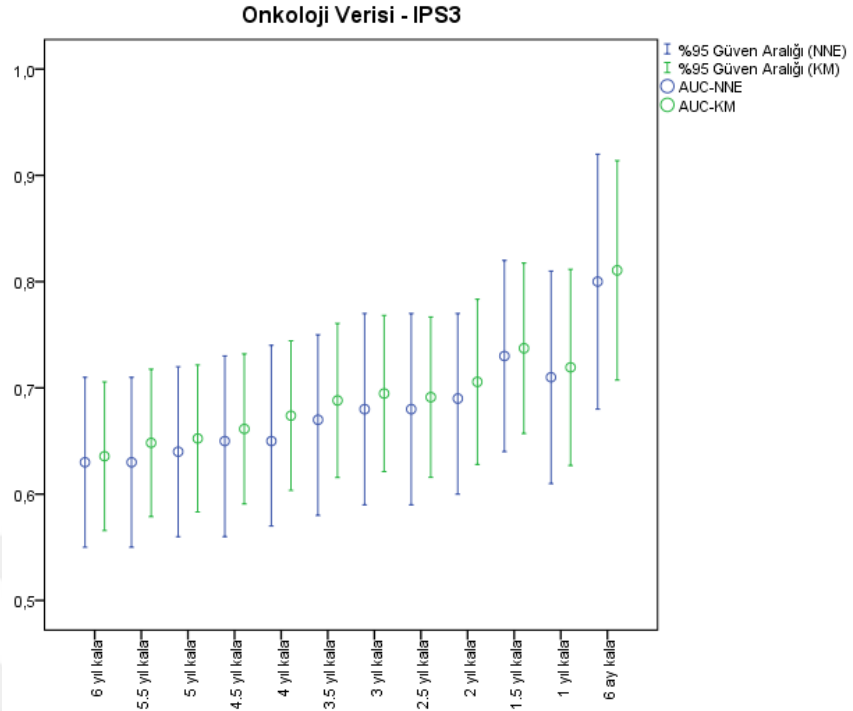
4.2.4.2. IPS3 Skoru Belirteç Olarak Alındığında

Hodgkin Lenfoma veri seti üzerinde IPS3 skorları için KM ve NNE yaklaşımlarının karşılaştırılması sonucu elde edilen bulgular Çizelge 4.20’de verilmiştir. Zaman dilimi 6 yıldan başlayıp 6 ay aralıklarla son 6 aya kadar devam etmektedir. Tüm zaman kesim noktalarında her iki yaklaşımdan elde edilen sonuçlar birbirine çok benzerdir ve %95 güven aralıklarından yola çıkılarak aralarında fark olmadığı söylenebilir.

Çizelge 4.20. Hodgkin Lenfoma veri setindeki IPS3 skoru için 6 yıl-6 ay zaman aralığında KM ve NNE yaklaşımları için elde edilen ROC eğrilerine ait EAKA değerleri ve %95 güven aralıkları

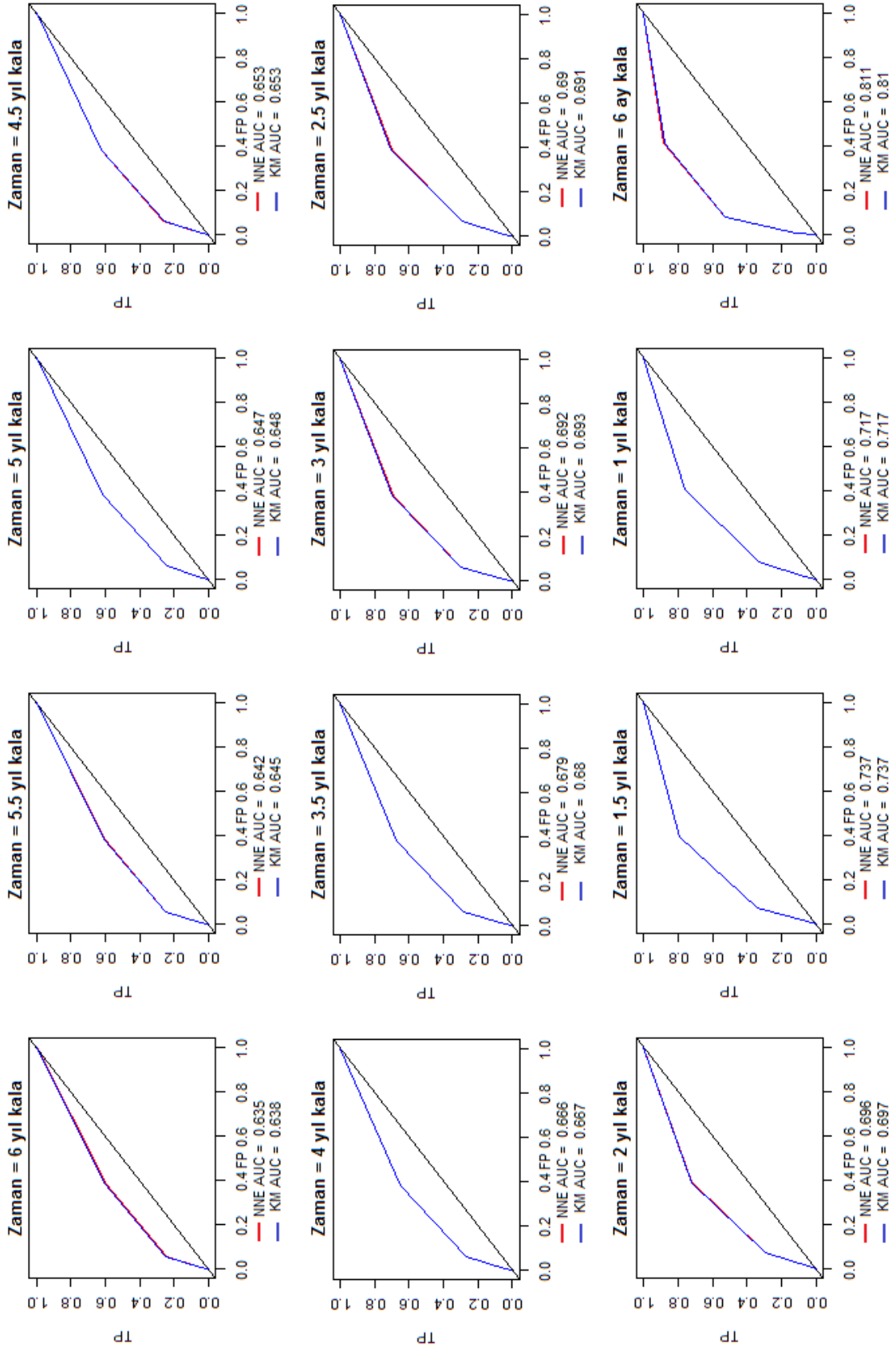
| Zaman | KM-EAKA | %95 GA | NNE-EAKA | %95 GA |
|--------------|---------|-------------|----------|-------------|
| 6 yıl kala | 0.64 | 0.57 - 0.71 | 0.64 | 0.55 - 0.71 |
| 5.5 yıl kala | 0.65 | 0.58 - 0.72 | 0.64 | 0.55 - 0.71 |
| 5 yıl kala | 0.65 | 0.58 - 0.72 | 0.65 | 0.56 - 0.72 |
| 4.5 yıl kala | 0.65 | 0.59 - 0.73 | 0.65 | 0.56 - 0.73 |
| 4 yıl kala | 0.67 | 0.60 - 0.74 | 0.67 | 0.57 - 0.74 |
| 3.5 yıl kala | 0.68 | 0.62 - 0.76 | 0.68 | 0.58 - 0.75 |
| 3 yıl kala | 0.69 | 0.62 - 0.77 | 0.69 | 0.59 - 0.77 |
| 2.5 yıl kala | 0.69 | 0.62 - 0.77 | 0.69 | 0.59 - 0.77 |
| 2 yıl kala | 0.70 | 0.63 - 0.78 | 0.70 | 0.60 - 0.77 |
| 1.5 yıl kala | 0.74 | 0.66 - 0.82 | 0.74 | 0.64 - 0.82 |
| 1 yıl kala | 0.72 | 0.63 - 0.81 | 0.72 | 0.61 - 0.81 |
| 6 ay kala | 0.81 | 0.71 - 0.91 | 0.81 | 0.68 - 0.92 |

Şekil 4.24’te KM ve NNE yaklaşımları ile elde edilen EAKA tahmin değerlerine ait hata çubukları grafiği verilmiştir. Tüm zaman kesim noktalarında yaklaşımların güven aralıklarının birbirlerini içerdiği görülmektedir.



Şekil 4.24. Farklı zamanlardaki IPS3 skorları için KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri EAKA'larının hata çubukları grafiği

Şekil 4.25'te KM ve NNE yaklaşımları ile çizdirilen ROC eğrileri karşılaştırılmaktadır. Tahminler çok benzer olduğundan eğriler birbiriyle çakışmaktadır.



Şekil 4.25. Hodgkin Lenfoma verisinde bulunan IPS3 belirteci üzerinde Kaplan-Meier ve Nearest Neighbor kestirimlerinin değişik zaman noktalarında karşılaştırılması

5. TARTIŞMA

ROC analizi, tanı testlerinin değerlendirilmesinde yaygın şekilde kullanılan bir yöntemdir. ROC eğrisi altında kalan alan (EAKA) bir doğruluk ölçütü olup, bu alan değeri 1'e ne kadar yakınsa testin veya biyolojik belirtecin ayırım başarısı o kadar iyidir. ROC eğrisinin çizilebilmesi için biyolojik belirteç ölçümlerinin ve statü bilgisinin (hasta-sağlıklı, öldü-ölmedi vb.) mevcut olması yeterlidir. Ancak bazı durumlarda ROC analizinin standart hali çalışmanın amacına yönelik olmayabilir. İlgilenilen olay veya belirtecin zaman içinde değişime uğraması, diğer bir deyişle zamana bağlı olması söz konusu ise, klasik ROC analizi yerine kullanılmak üzere geliştirilen zamana bağlı ROC analizine başvurulmalıdır.

Yüksek değerleri bir olayı (hastalık, ölüm, nüks vb.) gösteren zamana bağlı bir belirteç için ölçüm zamanı dikkate alınmadan yapılan bir değerlendirmede, değer düşük olmasına rağmen, olayın gerçekleştiği durumlar gözlenebilir. Diğer bir ifadeyle, zamana bağlı ölçümlerde olay yaklaştıkça ölçümün artması beklenir. Dolayısıyla ölçüm zamanı olaydan uzak olan bir ölçümün değerinin düşük olması nedeniyle olayı belirleyememe ihtimali olabilir. Bu durumda zaman bilgisini dikkate almayan klasik ROC analizi, belirtecin olayı belirlemedeki performansını olduğundan daha düşük belirleyecektir. Analizlerde zamanın etkisini dikkate alacak ve dolayısıyla tüm bireylerin ölçüm zamanlarını standardize edecek yöntemler kullanılarak zamana bağlı belirteçlerin sınıflama performansları belirlenebilir.

Bu tez çalışmasında, zamana bağlı ROC analizi yöntemlerinden en sık kullanılanları (KM, NNE) hem klasik ROC analizi ile hem de birbirleri ile karşılaştırılmış ve performansları değerlendirilmiştir.

Zamana bağlı ROC analizi literatürde ilk olarak 1999 yılında Etzioni ve ark.^[1] tarafından ortaya atılmıştır. Çalışmalarında uygulama verisi olarak prostat kanseri hastalarından oluşan ve belirteç ölçümünün PSA (Prostate Specific Antigen) olduğu bir boylamsal veri seti kullanmışlardır. Yazarlar, zaman boyutunu ROC analizi kavramına entegre ederken iki yaklaşımdan faydalanmışlardır. Birincisi veriyi modellemek, diğeri ise doğrudan ROC eğrisini modellemektir.

Heagerty ve ark.^[6], çalışmalarında sansürlü sağkalım verilerinde kullanılmak üzere parametrik varsayım içermeyen zamana bağlı ROC tahmin yöntemleri öne sürmüşlerdir. Belirtecin ölçüm zamanının sabit olduğunu (başlangıç zamanı) varsaymışlar ve dolayısıyla ROC eğrisini yalnızca olayın gerçekleşme zamanının bir fonksiyonu olarak ifade etmişlerdir. Ölçüm bilgisinin başlangıç zamanında sabitlenmiş olması ve yalnızca hastalık durumunun veya ilgilenilen olayın zamana bağlı değişmesine izin verilmesi Heagerty ve ark.'nın^[6] çalışmasında sunulan yaklaşımların bir kısıtlamasıdır.

Bu tezde uygulamalarda verilen kolon kanseri veri setinde pozitif nodül sayısının mortaliteyi tahminindeki başarısını değerlendirmek için klasik ROC analizi ile EAKA değeri 0.68 (%95 GA=0.63-0.73) olarak elde edilmiştir. EAKA üzerinde zaman etkisini görebilmek için zamana bağlı ROC analizinde zaman noktaları 3 yıldan başlayarak 6 aylık aralıklarla seçilmiştir (3, 2.5, 2, 1.5, 1, 0.5 yıl). Ölüme 3 yıl kala yapılan ölçümlerden elde edilen zamana bağlı EAKA değerleri klasik ROC ile aynı iken, 2.5 yıl ve daha yakın zamanlardaki EAKA değerleri giderek artmıştır (sırasıyla 0.69, 0.69, 0.71, 0.74, 0.76). Yalnızca son 1 yıl ve son 6 ay ölçümlerinin EAKA değerleri klasik ROC eğrisinin EAKA değerinden farklı bulunmuştur. Zamana bağlı EAKA'lar kendi aralarında karşılaştırıldığında ise, son 6 aya ait ölçümlerin performansının son 1 yıl dışında diğer tüm zaman noktalarından farklı olduğu görülmüştür. Geri kalan ölçümlerin tamamının EAKA değerleri birbirlerinin güven aralıklarının içine düşmüştür.

Melano veri setinde tümör kalınlığının mortaliteyi tahmin etmede ne kadar başarılı olduğu araştırılmıştır. Klasik ROC eğrisinden EAKA değeri 0.75 ($GA_{\text{klasik}}=0.68-0.82$) olarak elde edilmiştir. Burada da zamana bağlı analizler için ölüme 6 yıl kala başlayarak birer yıl aralıklarla yapılan ölçümler seçilmiştir (6, 5, 4, 3, 2, 1 yıl). Son iki yıla ait, yani ölüme 2 ve 1 yıl kala yapılan ölçümlerin, ölenleri ayırt etmede klasik ROC'tan daha başarılı olduğu ve bu farkın olarak anlamlı olduğu görülmüştür. Kendi içlerinde yapılan karşılaştırmada 2 yıl kala ($EAKA_2=0.88$, $GA_2=0.82-0.95$), 5 ve 6 yıl kaladan; 1 yıl kala ($EAKA_1=0.91$, $GA_1=0.85-0.97$) ise 3, 4, 5 ve 6 yıl kaladan daha iyidir (sırasıyla $EAKA_3=0.81$, $GA_3=0.73-0.89$, $EAKA_4=0.80$, $GA_4=0.73-0.88$, $EAKA_5=0.77$, $GA_5=0.67-0.84$ ve $EAKA_6=0.73$, $GA_6=0.63-0.81$).

Paquid veri setinde DSST ve MMSE olmak üzere iki ayrı belirtecin Alzheimer başlangıcını tahmin etmedeki başarısı ile ilgilenilmiştir. Bu veride de zaman noktaları ölüme son 12 yıldan başlayacak şekilde birer yıl arayla seçilmiştir. DSST için Klasik ROC analizi sonucu EAKA değeri 0.69 ($GA=0.66-0.71$) olarak elde edilmiştir. Zamana bağlı analizlerden elde edilen EAKA değerleri 12 yıl kala 0.74'ten başlayıp olaya yaklaştıkça 0.91'e ulaşmıştır. Klasik ROC eğrisinin EAKA değerine ait güven aralığı hiçbirini içermediğinden tüm yıllardaki ölçümlerin sınıflama performansı klasik ROC analizinden daha iyidir. Kendi içlerinde yapılan karşılaştırmada 10 yıl ve daha yakın zamanların 12 yıldan; 7 yıl ve daha yakın zamanların 11 yıldan; son 2 ve son 1 yılın ise hepsinden daha iyi olduğu görülmüştür.

Buna benzer sonuçlar MMSE belirteci için de elde edilmiştir. Klasik ROC eğrisinin EAKA değeri 0.63 ($GA_{\text{klasik}}=0.60-0.66$) olarak bulunmuştur. Zamana bağlı analizlerin tamamı klasik ROC'tan daha iyi sonuç vermiştir. Son 1 yıl kala yapılan ölçümlere ait EAKA değerinin olan standart hatası, diğer zamanlara ait standart hatalar [0.02-0.04] aralığında seyrederken, 0.07 olarak elde edilmiştir. Bu değer bu ölçümlere ait güven aralığının diğerlerine göre daha geniş olmasına sebep olmuştur. Bu durumun, belirteç değerlerinin bazı hastalarda olaya yaklaştıkça beklendiği kadar yükselmiyor olmasından kaynaklandığı düşünülmektedir.

Hodgkin Lenfoma verisinde Hodgkin Lenfoma tanısında kullanılan iki belirteç olan IPS7 ve IPS3'ün mortalite tahmin başarısı incelenmiştir. Zaman noktaları ölüme son 6 yıldan başlayacak şekilde altışar ay arayla seçilmiştir. IPS7 için klasik ROC analizi ile EAKA değeri 0.62 ($GA_{\text{klasik}}=0.54-0.69$) olarak elde edilmiştir. Klasik ROC eğrisine ait güven aralığı, zamana bağlı çizilen ROC eğrilerinin EAKA değerlerinin son 6 ay hariç tamamını içermektedir. Yalnızca son 6 ay ölçümlerinin sınıflama başarısı klasik ROC'unkinden olarak yüksektir ($EAKA=0.75$). IPS3 için ise, klasik ROC ile elde edilen EAKA değeri 0.65 ($GA_{\text{klasik}}=0.58-0.72$) olmuştur. Ölüme 1.5 yıl ($EAKA=0.74$) ve 6 ay kala ($EAKA=0.81$) yapılan ölçümlerin tanı performansı klasik ROC'unkinden daha iyidir. Her iki belirteç için de 6 yıldan son 6 aya kadar elde edilen EAKA değerlerinin monoton artmadığı, bazı noktalarda artıp azaldığı gözlenmiştir. Bu durumun skorların değer aralığının çok dar olmasından kaynaklandığı düşünülmektedir.

Literatüre benzer şekilde, belirteçlerin farklı zaman noktalarındaki ölçümlerinin karşılaştırılması sonucunda olaya yaklaştıkça EAKA tahmin değerlerinin (bazı istisnai durumlar dışında) arttığı görülmüştür. Martinez-Cambor ve ark.'nın^[22] çalışmasındaki gerçek veri uygulamasında, takip süresinin başında elde edilen ölçümlerin EAKA değeri zaman içinde 0.73'ten 0.68'e geriledikten sonra yeniden artmaya başlamış ve 14 yıl içinde (olaya yaklaştıkça) 0.76'ya yükselmiştir. Bu dalgalanmaların uzun takip aralıklarında belirteçlerin tanısal değerinin kaybolmasından kaynaklandığı düşünülmektedir.

Literatürde klasik ROC analizi ile zamana bağlı ROC analizi karşılaştırmasına yer veren yeterli sayıda çalışma bulunmamakla birlikte, Chambless ve Diao'nun^[38] çalışması bu doğrultudaki az sayıda eserden biridir. Yazarlar, tahminlerinde zamanı göz önünde bulunduran bir EAKA yöntemi ile zaman etkisini göz ardı eden bir yöntemin aynı sonucu veremeyeceğini ifade etmiştir. Bu tez çalışmasındaki uygulamalardan elde edilen bulgular bunu destekler niteliktedir. Ancak yazarlar risk faktörlerinin lineer bir kombinasyonunun incelendiği durumda, NNE yaklaşımının, faktörlerin tek tek katsayılarını hesaba katmayacağını belirtmiştir. Ayrıca KM yaklaşımında sağkalım fonksiyonu ($S(t|z)$) için bir kestirim elde edilmemesinin, yöntemin bir kısıtlaması olduğunu ifade etmiştir.

Heagerty ve ark.^[6], Heagerty ve Zheng^[2], Cai ve ark.^[9], Song ve Zhou'nun^[39] yaptığı çalışmalar genellikle farklı belirteçlerin performanslarının zamana bağlı ROC analizi kullanılarak karşılaştırılması üzerinedir. Bu çalışmada da benzer şekilde, Paquid verisindeki DSST ve MMSE belirteçleri hem KM hem NNE yöntemi kullanılarak karşılaştırılmış ve her iki yöntemde de DSST'nin ayırım başarısı daha yüksek bulunmuştur.

Blanche ve ark.^[40] değişik sansür durumlarında kullanılan ve KM ile NNE'nin de aralarında bulunduğu parametrik olmayan zamana bağlı ROC eğrisi kestirimlerini incelemiştir. Yazarlar, sansürün belirtece bağlı olduğu verilerde KM kestiriminin NNE'ye göre yanlış olduğu ve bu yanlışlığın belirteç ve sansür zamanı arasındaki ilişkiyle birlikte arttığı sonucuna varmıştır. Diğer taraftan, bazı komşuluklarda ilgilenilen zamanda risk altındaki birey sayısı sıfır olduğunda NNE yaklaşımının sorun yaşayabileceği belirtilmektedir. Ancak bu problem yalnızca küçük örneklerde, çok sayıda sansür

varlığında ve belirteç ile sansürleme arasındaki ilişkinin yüksek olduğu durumda ortaya çıkabileceğinden, önüne geçilmesi mümkündür.

Literatüre paralel olarak bu çalışmada kullanılan tüm veri setlerinde, t zaman noktalarındaki KM ve NNE kestirimleri karşılaştırıldığında sonuçların benzer olduğu görülmüştür^[6,40]. KM genellikle daha yüksek alan değerleri verdiği halde, sansür sürecinin biyolojik belirtece bağlı olmasına izin vermemek gibi bir kısıtlamaya sahiptir. Buna karşılık NNE için böyle bir kısıtlama söz konusu değildir. Hesaplanan alan değerleri daha düşük olsa da, NNE yaklaşımı ile KM'ye göre daha düzgünleştirilmiş eğriler elde edilmiştir. Bu sayede ROC eğrisinin monoton artan fonksiyon olma özelliği sağlanmaktadır.

Bu çalışmada incelenen zamana bağlı ROC analizi yöntemlerinin (KM, NNE) temelinde, Etzioni ve ark.'nın^[1] önerdiği “verinin modellenmesi” yaklaşımı bulunmaktadır. Verinin modellenmesi, mevcut zamanlardaki ölçümler kullanılarak, belirteç bilgisi bulunmayan zaman noktalarına dair ölçüm verisinin modellenmesi anlamına gelmektedir. Bu model doğru olsa bile, tahmin değerlerinin varyansları gerçek varyansları olduğundan düşük kestirecektir. Bu nedenle, ilgilenilen t zamanındaki gözlemleri tahmin etmek ve bu tahminleri kullanarak ROC eğrisini çizdirmek; bir yapay iyimserlik eğiliminde olacağından, tanı testinin veya belirtecin başarısını belirlemede yanlış davranacaktır.

6. SONUÇ VE ÖNERİLER

Bu tez çalışmasında, zamanın sabit bir özellik olduğu klasik ROC analizinden uyarlanmış bir yöntem olan zamana bağlı ROC analizi hakkında bilgi verilmiştir. Teorik açıdan ele alınan yöntem, çeşitli uygulamalar yoluyla desteklenmiştir.

Bir zamana bağlı ROC analizi yaklaşımı olan Kaplan-Meier kestirimi, sağkalım verilerinde olayın gerçekleşmesini tahmin etmede klasik ROC analizine göre daha iyi sonuçlar vermiştir. Ayrıca, farklı zaman noktalarında yapılmış ölçümlerden elde edilen zamana bağlı ROC eğrileri, ilgilenilen olaya yaklaştıkça biyolojik belirtecin olayı tahmin etme yeteneğinin arttığını göstermiştir. Böylece ölçümün yapıldığı zamanın da tanıda ölçüm değerleri kadar büyük önem taşıdığı doğrulanmıştır. Eğer ilgilenilen olay veya belirteç zaman içinde değişkenlik gösteriyorsa, yani zamana bağlılık söz konusu ise, tanı testinin performansını ölçmede klasik ROC analizi yerine zamanı göz önünde bulunduran zamana bağlı ROC analizi kullanılmalıdır.

İki yöntemin karşılaştırılmasından elde edilen bulgular, Kaplan-Meier ve En Yakın Komşu kestirimlerinin birbirine yakın sonuçlar verdiğini göstermiştir. Ancak ikisinden birinin tercih edilmesi gerektiğinde, ROC eğrisinin monoton artan fonksiyon olması özelliğini garanti etmesi ve sansür sürecinin biyolojik belirtece bağlı olmasına izin vermesi nedeniyle En Yakın Komşu kestiriminin seçilmesi daha uygun olacaktır.

Bundan sonra yapılacak çalışmalarda, veriyi değil ROC eğrisini modellemeye yönelik yöntemler incelenebilir veya zamana bağlı ROC analizinde optimal kesim noktaları belirlenmesi için yöntemlerin geliştirilmesi ele alınabilir.

KAYNAKLAR

1. **Etzioni R, Pepe MS, Longton G, Chengcheng H, Goodman G.** Incorporating the time dimension in receiver operating characteristic curves: a case study of prostate cancer. *Medical Decision Making*, **1999**; 19:242-251.
2. **Heagerty PJ, Zheng Y.** Survival model predictive accuracy and ROC curves. *Biometrics*, **2005**; 61:92-105.
3. **Zheng Y, Heagerty PJ.** Prospective accuracy for longitudinal markers. *Biometrics*, **2007**; 63:332-341.
4. **Zheng Y, Heagerty PJ.** Semiparametric estimation of time-dependent ROC curves for longitudinal marker data. *Biostatistics*, **2004**; 5(4):615-632.
5. **Antolini L, Boracchi P, Biganzoli E.** A time-dependent discrimination index for survival data. *Statistics in Medicine*, **2005**; 24:3927-3944.
6. **Heagerty PJ, Lumley T, Pepe MS.** Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*, **2000**; 56:337-344.
7. **Akritas M.** Nearest neighbor estimation of a bivariate distribution under random censoring. *Annals of Statistics*, **1994**; 22:1299-1327.
8. **Slate E, Turnbull B.** Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine*, **2000**; 19:617-637.
9. **Cai T, Pepe MS, Zheng Y, Lumley T, Jenny N.** The sensitivity and specificity of markers for event times. *Bioinformatics*, **2006**; 7(2):182-197.
10. **Leisenring W, Pepe MS, Longton G.** A marginal regression modelling framework for evaluating medical diagnostic tests. *Statistics in Medicine*, **1997**; 16:1263-1281.
11. **Rizopoulos D.** Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, **2011**; 67:819-829.
12. **Wulfsohn M, Tsiatis AA.** A joint model for survival and longitudinal data measured with error. *Biometrics*, **1997**; 53:330-339.
13. **Henderson R, Diggle P, Dobson A.** Joint modelling of longitudinal measurements and event time data. *Biostatistics*, **2000**; 1(4):465-480.

14. **Tsiatis AA, Davidian M.** Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, **2004**; 14(3):809-834.
15. **Pepe MS.** A regression modeling framework for ROC curves in medical diagnostic testing. *Biometrika*, **1997**; 84:595-608.
16. **Cai T, Pepe MS.** Semiparametric receiver operating characteristic analysis to evaluate biomarkers for disease. *Journal of the American Statistical Association*, **2002**; 97:1099-1107.
17. **Zheng Y, Cai T, Feng Z.** Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. *Biometrics*, **2006**; 62:279-287.
18. **Li J, Ma S.** Time-dependent ROC analysis under diverse censoring patterns. *Statistics in Medicine*, **2011**; 30:1266-1277.
19. **Hung H, Chiang CT.** Optimal composite markers for time-dependent receiver operating characteristic curves with censored survival data. *Scandinavian Journal of Statistics*, **2010**; 37:664-679.
20. **Liu D, Cai T, & Zheng Y.** Evaluating the predictive value of biomarkers with stratified case-cohort design. *Biometrics*, **2012**; 68:1219-1227.
21. **Li S, Ning Y.** Estimation of covariate-specific time-dependent ROC curves in the presence of missing biomarkers. *Biometrics*, **2015**; 71:666-676.
22. **Martinez-Cambor P, Bayon GF, Perez-Fernandez S.** Cumulative/dynamic ROC curve estimation. *Journal of Statistical Computing and Simulation*. **2016**; 86(17), 3582-3594, DOI: 10.1080/00949655.2016.1175442.
23. **Karaismailođlu E.** İzlem süresi içerisinde belirteçlerin performanslarını değerlendirilmede zamana bađlı ROC eğrisinin kullanımı. Doktora tezi, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara, **2015**.
24. **Pepe MS.** The Statistical Evaluation of Medical Tests for Classification and Prediction. New York: Oxford University Press, **2003**.
25. **Karaađaođlu E, Karakaya J, Kılıçkap M.** Tanı Testlerinin Deđerlendirilmesinde İstatistiksel Yöntemler. Ankara: Detay Yayıncılık, **2016**.
26. **Ünal, İ.** Tanı Testlerinin Deđerlendirilmesinde Kullanılan Standartlar ve Analitik Yöntemler. Doktora tezi, Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü, Adana, **2010**.
27. **Zhou XH, Obuchowski NA, McClish DK.** Statistical Methods in Diagnostic Medicine. New York: Wiley, **2002**.

28. **Fleiss J, Levin B, Paik MC.** Statistical Methods for Rates and Proportions, 3rd Ed, New Jersey: Wiley, **2003**.
29. **Tosteson AN, Begg CB.** A general regression methodology for ROC curve estimation. *Medical Decision Making*, **1988**; 3:204-215.
30. **Pepe MS.** Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. *Biometrics*, **1998**; 54:124-35.
31. **Balasubramanian R, Lagakos S.** Estimation of the timing of perinatal transmission of HIV. *Biometrics*, **2001**; 57:1048-1058.
32. **Kaplan E, Meier P.** Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **1958**, 53:457-481.
33. **Blanche P, Dartigues JF, Jacqmin-Gadda, H.** Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in Medicine*, 32(30), 5381-5397. URL <http://onlinelibrary.wiley.com/doi/10.1002/sim.5958/full>, **2013**.
34. **R Development Core Team.** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>, **2015**.
35. **IBM Corp.** IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp, **2011**.
36. **Heagerty PJ and packaging by Paramita Saha-Chaudhuri.** survivalROC: Time-dependent ROC curve estimation from censored survival data. R package version 1.0.3. <http://CRAN.R-project.org/package=survivalROC>, **2013**.
37. **Diefenbach CS, Li H, Hong F, Gordon LI, Fisher RI, Bartlett NL, ..., Advani RH.** Evaluation of the international prognostic score (IPS-7) and a simpler prognostic score (IPS-3) for advanced hodgkin lymphoma in the modern era. *Br J Haematol*, **2015**; 171(4):530-538.
38. **Chambless L, Diao G.** Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Statistics in Medicine*, **2006**; 25:3474-3486.
39. **Song X, Zhou XH.** A semiparametric approach for the covariate specific ROC curve with survival outcome. *Statistica Sinica*, **2008**; 18:947-965.
40. **Blanche P, Dartigues JF, Jacqmin-Gadda, H.** Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring. *Biometrical Journal*, **2013**; 5:687-704.

EKLER

Analizlerde Kullanılan R Kodları

```
library(survival)
library(timeROC)
library(survivalROC)
library(bootstrap)

###NNE Güven Aralıkları için Fonksiyon

survROC<-function(x,veri)
{
  zaman<-veri$zaman[x]
  sansur<-veri$sansur[x]
  belirtec<-veri$belirtec[x]
  kesimnoktasi<-veri$kesimnoktasi
  nobs<-veri$nobs

  teta0<-survivalROC(Stime=zaman, status=sansur, marker=belirtec, predict.time=kesimnoktasi,
span=0.25*(nobs)^(-0.20))

  return(teta0$AUC)
}

perc95 <- function(x){quantile(x, .5)}
results <- bootstrap(1:n,1000,survROC,veri,func = perc95)      #n örneklem büyüklüğü
mn<-mean(results$thetastar)
ss<-sd(results$thetastar)
round(c(mn-1.96*ss,mn,mn+1.96*ss),4)
```

```
###Kolon Kanseri Verisi için Kodlar
```

```
veri500 <- read.table("C:/Users/Ceren/Desktop/TEZ/uygulamalar/kolon kanseri verisi - kesilmiş/veri500.txt",header=T)
```

```
#survivalROC
```

```
nobs <- NROW(veri500)
```

```
kesimnoktasi <- 365*3 #Farklı zaman kesim noktaları için değiştirilebilir.
```

```
nod_nne = survivalROC(Stime=veri500$stime,
```

```
sansur=veri500$sansur,
```

```
belirtec = veri500$pnod,
```

```
predict.time = kesimnoktasi,span = 0.25*nobs^(-0.20) )
```

```
nod_km = survivalROC(Stime=veri500$stime,
```

```
sansur=veri500$sansur,
```

```
belirtec = veri500$pnod,
```

```
predict.time = kesimnoktasi, method="KM")
```

```
plot(nod_nne$FP, nod_nne$TP, type="l", xlim=c(0,1), ylim=c(0,1),
```

```
xlab=paste( "FP", "\n", "NNE AUC = ",round(nod_nne$AUC,3),
```

```
"\n", "KM AUC = ", round(nod_km$AUC,3)),
```

```
ylab="TP", main="Zaman = 3 yıl kala", col="red")
```

```
abline(0,1)
```

```
lines(nod_km$FP, nod_km$TP, type="l", xlim=c(0,1), ylim=c(0,1), col="blue")
```

```
#timeROC
```

```
ROC.pnod.marginal<-timeROC(T=veri500$stime,
```

```
delta=veri500$sansur,belirtec=veri500$pnod,
```

```
cause=1,weighting="marginal",
```

```
times=quantile(veri500$stime,probs=seq(0.2,0.8,0.1)), #İstenen zaman noktaları yazılabilir
```

```
iid=TRUE)
```

```
ROC.pnod.marginal
```

```
###Melano Verisi için Kodlar
```

```
mel <- read.table("C:/Users/Ceren/Desktop/TEZ/uygulamalar/verilerin R hali/melano data - kes.txt",header=T)
```

```
#survivalROC
```

```
nobs <- NROW(mel)
```

```
kesimnoktasi <- 365 #Farklı zaman kesim noktaları için değiştirilebilir.
```

```
tumor_nne = survivalROC(Stime=mel$time,
```

```
sansur=mel$sansur,
```

```
belirtec = mel$thick,
```

```
predict.time = kesimnoktasi,span = 0.25*nobs^(-0.20) )
```

```
tumor_km = survivalROC(Stime=mel$time,
```

```
sansur=mel$sansur,
```

```
belirtec = mel$thick,
```

```
predict.time = kesimnoktasi, method="KM")
```

```
plot(tumor_nne$FP, tumor_nne$TP, type="l", xlim=c(0,1), ylim=c(0,1),
```

```
xlab=paste( "FP", "\n", "NNE AUC = ",round(tumor_nne$AUC,3),
```

```
"\n" , "KM AUC = ", round(tumor_km$AUC,3)),
```

```
ylab="TP", main="Zaman = 1 yıl kala", col="red")
```

```
abline(0,1)
```

```
lines(tumor_km$FP, tumor_km$TP, type="l", xlim=c(0,1), ylim=c(0,1), col="blue")
```

```
#timeROC
```

```
ROC.thick.marginal<-timeROC(T=mel$time,
```

```
delta=mel$sansur,belirtec=mel$thick,
```

```
cause=1,weighting="marginal",
```

```
times=quantile(mel$time,probs=seq(0.2,0.8,0.1)), #İstenen zaman noktaları yazılabilir
```

```
iid=TRUE)
```

```
ROC.thick.marginal
```

```
###Paquid Verisi için Kodlar (DSST ve MMSE için aynı)
```

```
paq <- read.table("C:/Users/Ceren/Desktop/TEZ/uygulamalar/verilerin R hali/paquid data - kes.txt",header=T)
```

```
#survivalROC
```

```
nobs <- NROW(paq)
```

```
kesimnoktasi <- 1 #Farklı zaman kesim noktaları için değiştirilebilir.
```

```
digit_nne = survivalROC(Stime=paq$time,
```

```
sansur=paq$sansur,
```

```
belirtec = -paq$DSST,
```

```
predict.time = kesimnoktasi,span = 0.25*nobs^(-0.20) )
```

```
digit_km = survivalROC(Stime=paq$time,
```

```
sansur=paq$sansur,
```

```
belirtec = -paq$DSST,
```

```
predict.time = kesimnoktasi, method="KM")
```

```
plot(digit_nne$FP, digit_nne$TP, type="l", xlim=c(0,1), ylim=c(0,1),
```

```
xlab=paste( "FP", "\n", "NNE AUC = ",round(digit_nne$AUC,3),
```

```
"\n" , "KM AUC = ", round(digit_km$AUC,3)),
```

```
ylab="TP", main="NNE-KM \n Zaman = 1 yıl", col="red")
```

```
abline(0,1)
```

```
lines(digit_km$FP, digit_km$TP, type="l", xlim=c(0,1), ylim=c(0,1), col="blue")
```

```
#timeROC
```

```
ROC.DSST<-timeROC(T=paq$time,
```

```
delta=paq$sansur,belirtec=paq$DSST,
```

```
cause=1,weighting="marginal",
```

```
times=quantile(paq$time,probs=seq(0.2,0.8,0.1)), #İstenen zaman noktaları yazılabilir
```

```
iid=TRUE)
```

```
ROC.DSST
```

```
###Hodgkin Lenfoma Verisi için Kodlar (IPS7 ve IPS3 için aynı)
```

```
onko <- read.table("C:/Users/Ceren/Desktop/TEZ/uygulamalar/verilerin R  
hali/onko_r.txt",header=T)
```

```
#survivalROC
```

```
nobs <- NROW(onko)
```

```
kesimnoktasi <- 2*365
```

```
hlenfoma8_nne = survivalROC(Stime=onko$time_os,
```

```
sansur=onko$sansur,
```

```
belirtec = onko$ips7,
```

```
predict.time = kesimnoktasi,span = 0.25*nobs^(-0.20) )
```

```
hlenfoma8_km = survivalROC(Stime=onko$time_os,
```

```
sansur=onko$sansur,
```

```
belirtec = onko$ips7,
```

```
predict.time = kesimnoktasi, method="KM")
```

```
plot(hlenfoma8_nne$FP, hlenfoma8_nne$TP, type="l", xlim=c(0,1), ylim=c(0,1),
```

```
xlab=paste( "FP", "\n", "NNE AUC = ",round(hlenfoma8_nne$AUC,3),
```

```
"\n" , "KM AUC = ", round(hlenfoma8_km$AUC,3)),
```

```
ylab="TP", main="Zaman = 2 yıl kala", col="red")
```

```
abline(0,1)
```

```
lines(hlenfoma8_km$FP, hlenfoma8_km$TP, type="l", xlim=c(0,1), ylim=c(0,1), col="blue")
```

```
#timeROC
```

```
ROC.IPS7.survival<-timeROC(T=onko$time_os,
```

```
delta=onko$sansur,belirtec=onko$ips7,
```

```
cause=1,weighting="marginal",
```

```
times=c(0.5*365,365,1.5*365,2*365,2.5*365),iid=TRUE)
```

```
ROC.IPS7.survival
```

ÖZGEÇMİŞ

1991 yılında İzmir’de doğdu. İlkokulu ve ortaokulu Kahramanmaraş Atatürk İlköğretim Okulu’nda (1997-2005), liseyi Kahramanmaraş Çukurova Elektrik Anadolu Lisesi’nde (2005-2009) okudu. Üniversite öğrenimini Orta Doğu Teknik Üniversitesi İstatistik Bölümü’nde tamamladı (2009-2014). 2014-2015 öğretim yılı güz döneminde, Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı’nda Yüksek Lisans programına kabul edildi. 2015 yılının Ocak ayında başvurduğu Öğretim Üyesi Yetiştirme Programı (ÖYP) ile Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı’nda araştırma görevlisi kadrosuna yerleşti. O tarihten bu yana Kahramanmaraş Sütçü İmam Üniversitesi adına ÖYP kapsamında Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı’nda hem Yüksek Lisans öğrenimi görmekte hem de araştırma görevlisi olarak görev yapmaktadır.