**T.C.**
**YILDIZ TEKNİK ÜNİVERSİTESİ**
**FEN BİLİMLERİ ENSTİTÜSÜ**

# PRENATAL TANIDA TRİZOMİ RİSK TAHMİNİ İÇİN QF-PCR SONUÇLARI, DEMOGRAFİK, KLİNİK VE GENETİK VERİLERİN ENTEGRASYONUNUN MAKİNE ÖĞRENMESİ TABANLI ANALİZİ

**Kübra TÜREGÜN**

YÜKSEK LİSANS TEZİ
Biyoteknoloji Anabilim Dalı
Biyoteknoloji Programı

Danışman
Doç. Dr. Alper YILMAZ

Şubat, 2025

**T.C.**
**YILDIZ TEKNİK ÜNİVERSİTESİ**
**FEN BİLİMLERİ ENSTİTÜSÜ**

# PRENATAL TANIDA TRİZOMİ RİSK TAHMİNİ İÇİN QF-PCR SONUÇLARI, DEMOGRAFİK, KLİNİK VE GENETİK VERİLERİN ENTEGRASYONUNUN MAKİNE ÖĞRENMESİ TABANLI ANALİZİ

Kübra TÜREGÜN tarafından hazırlanan tez çalışması 20.02.2025 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Biyoteknoloji Anabilim Dalı Biyoteknoloji Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Alper YILMAZ
Yıldız Teknik Üniveristesi
Danışman

**Jüri Üyeleri**

Doç. Dr. Alper YILMAZ, Danışman
Yıldız Teknik Üniveristesi _____

Prof. Dr. Sevgi MARAKLI, Üye
Yıldız Teknik Üniversitesi _____

Prof. Dr. Şükrü ÖZTÜRK, Üye
İstanbul Üniversitesi _____

Danışmanım Doç. Dr. Alper YILMAZ sorumluluğunda tarafımca hazırlanan "Prenatal Tanıda Trizomi Risk Tahmini için QF-PCR Sonuçları, Demografik, Klinik ve Genetik Verilerin Entegrasyonunun Makine Öğrenmesi Tabanlı Analizi" başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Kübra TÜREGÜN

İmza

*Canım Anneme*
*ve Sevgili Ceviz'ime*

# TEŞEKKÜR

Şüphesiz, bazı insanlar hayatımızı zenginleştirir ve hedeflerimize ilerlerken emeklerimize anlam katar. Tez çalışmam sürecinde, böyle değerli insanlarla çevrili olduğum için kendimi şanslı hissediyorum. Bu akademik yolculuğumda, değerli tez danışmanım Doç. Dr. Alper Yılmaz'a sonsuz teşekkürlerimi sunmak istiyorum. Kendisi, bir danışmanın ötesinde, araştırma tutkusunu paylaşarak beni motive eden, eleştirel düşünme becerilerimi geliştirmemde rehberlik eden bir mentor oldu. Bilimsel araştırmanın inceliklerini sabırla aktaran ve her aşamada desteğini esirgemeyen değerli hocama en içten şükranlarımı sunarım. Akademik ve profesyonel yaşamımın en kritik dönemlerinden birinde, Doç. Dr. Alper Gezdirici'nin katkıları benim için büyük bir anlam taşımaktadır. Zor ve çetrefilli bir süreçte hastane süreçlerinde gösterdiği profesyonel yaklaşımı ve desteğiyle benim için hayati bir katkı sağladı. İnsan odaklı yaklaşımı ve sarsılmaz desteği sayesinde, en zorlu günlerimi aşabilme gücünü buldum. Kendisine en derin şükranlarımı ve sonsuz minnettarlığımı ifade etmek istiyorum. Çalışma sürecimde katkı ve desteklerinden dolayı İstanbul Başakşehir Çam ve Sakura Şehir Hastanesi'ne içtenlikle teşekkür ederim. Hayatımın her döneminde maddi ve manevi desteklerini hiçbir zaman esirgemeyen, eğitimimi her şeyin üstünde tutan sevgili annem ve babam Meşhur ve Ramazan Türegün'e sonsuz minnettarlığımı sunarım. Onların özverisi ve koşulsuz sevgisi, bugünlere gelmemde en büyük güç kaynağım olmuştur. Yaşamımın her döneminde olduğu gibi, bu süreçte de desteklerini ve sevgilerini daima hissettiğim, moral kaynağım olan sevgili kardeşlerim Gülten, Esma ve Mehmet Türegün'e, teknik destekçilerim ve ev arkadaşlarım Hasan ve Enes'e, minik kalbi ile bana şans dileyen Eyüp'e; her an bir telefon uzağımda olan, desteğini ve dostluğunu hiç eksik etmeyen değerli dostum, avukatım, kardeşim Medine Tufan'a; varlığı ve desteği bu yolculukta bana güç veren, bu çalışmanın baş mimarlarından, sadece bir dost değil, aynı zamanda yol göstericim olan Nuri Acar'a yürekten teşekkür ederim. Zamanını ve emeğini koşulsuzca adayan, tez çalışmamın her aşamasında sarsılmaz inancı ve sonsuz sabrıyla yanımda olan, beni benden daha iyi anlayan değerli yol arkadaşıma tüm katkıları için minnettarım. Varlıkları ve anlayışları bu süreçte bana güç vermiştir.

Kübra TÜREGÜN

# İÇİNDEKİLER

**KISALTMA LİSTESİ** viii

**ŞEKİL LİSTESİ** x

**TABLO LİSTESİ** xii

**ÖZET** xiii

**ABSTRACT** xv

**1 GİRİŞ** 1

1.1 Literatür Özeti . . . . . 1
1.2 Tezin Amacı . . . . . 2
1.3 Hipotez . . . . . 3

**2 PRENATAL TANI VE PRENATAL TANI TESTLERİ** 4

2.1 Prenatal Tanı . . . . . 4
2.2 Gebelik Dönemleri ve Özellikleri . . . . . 5
2.2.1 Birinci Trimester (1-13. Hafta) . . . . . 5
2.2.2 İkinci Trimester (14-26. Hafta) . . . . . 6
2.2.3 Üçüncü Trimester (27-40. Hafta) . . . . . 6
2.3 Prenatal Tanı Testleri . . . . . 6
2.3.1 Non-İnvaziv Testler . . . . . 8
2.3.2 İnvaziv Testler . . . . . 9

**3 KROMOZOMAL ANOMALİLERİN TANISINDA QF-PCR VE DİĞER YÖNTEMLER** 12

3.1 Kromozomal Anöploidi Tanısında QF-PCR Metodolojisi . . . . . 13
3.2 Kromozom Anomalileri . . . . . 17
3.2.1 Anöploidiler . . . . . 17
3.3 Kromozomal Anöploidi Tespitinde Diğer Tanısal Yöntemler . . . . 25
3.3.1 Array-CGH . . . . . 26
3.3.2 Kromozomal Mikroarray Analizi (CMA) . . . . . 26
3.3.3 Karyotipleme . . . . . 27

3.3.4 FISH-Floresan In Situ Hibridizasyon . . . . . 28

**4 MAKİNE ÖĞRENMESİ ALGORİTMALARI 29**

4.1 Makine Öğrenmesi Algoritmaları . . . . . 29

4.1.1 Lojistik Regresyon (Logistic Regression) Algoritması . . . 29

4.1.2 Karar Ağacı (Decision Tree) Algoritması . . . . . 30

4.1.3 Random Forest (Rastgele Orman) Algoritması . . . . . 30

4.1.4 K-Nearest Neighbor (KNN, K-En Yakın Komşu) Algoritması 31

4.2 Prenatal Tanıda Kullanılan Makine Öğrenmesi ve Yapay Zeka Uygulamaları . . . . . 32

**5 MATERYAL-METOT 34**

5.1 Veri Kaynakları ve Toplanma Süreci . . . . . 34

5.1.1 Çalışma Popülasyonunun Sınıflandırılması . . . . . 34

5.2 Veri Yönetimi ve Ön İşleme . . . . . 37

5.2.1 Veri Doğrulama Süreci ve Kullanılan R Kodu . . . . . 37

5.2.2 Veri Temizleme ve Standardizasyon . . . . . 38

5.3 Verilerin Sistematik Sınıflandırması . . . . . 43

5.4 Çalışma Ortamı ve Kullanılan R Paketleri . . . . . 44

5.5 İstatistiksel Analizler . . . . . 46

5.6 Dağılımlara Yönelik Tamamlayıcı Analizler . . . . . 47

5.6.1 Yaş Verisinin Görselleştirilmesi ve QF-PCR Sonuçlarıyla İlişkisi . . . . . 47

5.6.2 USG Bulguları ve QF-PCR Sonuçlarının Tanımlayıcı ve Görselleştirilmiş Analizi . . . . . 48

5.6.3 Prenatal Tarama Testlerinin Terminasyon Kararlarına Etkisinin Dağılım Analizi . . . . . 48

5.7 Makine Öğrenmesi Entegrasyonu . . . . . 49

**6 TARTIŞMA 53**

6.1 Çalışmanın Klinik Önemi ve Literatür Karşılaştırması . . . . . 53

6.2 Metodolojik Değerlendirme ve Model Performansı . . . . . 55

6.3 Kısıtlılıklar ve Gelecek Perspektifi . . . . . 56

**7 SONUÇ 58**

7.1 Demografik, Klinik ve Genetik Verilerin Analizi . . . . . 58

7.1.1 Popülasyon Analizi ve Yaş Dağılımı . . . . . 58

7.1.2 Klinik Bulgular ve Gebelik Özellikleri . . . . . 59

7.1.3 Genetik Analiz Sonuçları ve Yaş İlişkisi . . . . . 59

7.1.4 Prenatal Tanı Endikasyonları ve Numune Dağılımı . . . . . 61

7.1.5 QF-PCR Sonuçları ve Gebelik Sonuçları İlişkisi . . . . . 61
7.1.6 Maternal ve Paternal Kalıtım Analizi . . . . . 61
7.2 İstatistiksel Analiz Sonuçları . . . . . 63
7.3 Makine Öğrenmesi Modellerinin Sonuçları . . . . . 66
7.3.1 Veri Hazırlama ve Normalizasyon . . . . . 66
7.3.2 Kategorik Değişkenlerin Kodlanması . . . . . 66
7.3.3 Eğitim ve Test Veri Setleri . . . . . 67
7.3.4 Makine Öğrenmesi Modellerinin Karşılaştırmalı Performans Analizi . . . . . 67
7.3.5 Değişken Önem Dereceleri ve Model Performansı . . . . . 71
7.3.6 Karar Ağacı Modelinin Analizi . . . . . 73
7.4 Prenatal Tarama Testleri ve Terminasyon Kararlarının Değerlendirilmesi . . . . . 76
7.5 Prenatal Tanı Süreçlerinde Makine Öğrenmesi Modellerinin Entegrasyonu ve Trizomi Risk Tahmini . . . . . 78

**KAYNAKÇA** **81**

**TEZDEN ÜRETİLMİŞ YAYINLAR** **89**

# KISALTMA LİSTESİ

| | |
|---|---|
| AFP | Alfa-fetoprotein |
| AI-YZ | Artificial Intelligence (Yapay Zeka) |
| ANN | Artificial Neural Network (Yapay Sinir Ağı) |
| Array-CGH | Array Comparative Genomic Hybridization (Mikrodizi Karşılaştırmalı Genomik Hibridizasyon) |
| ARSA | Aberran Sağ Subklavian Arter |
| ASD | Atriyal Septal Defekt |
| AST | Alternate Step Test |
| AUC | Area Under Curve (ROC Eğrisinin Altında Kalan Alan) |
| AVSD | Atriyoventriküler Septal Defekt |
| Bp-bç | Base pair (Baz çifti) |
| $\beta$-hCG | Beta İnsan Koryonik Gonadotropini |
| cffDNA | Cell-free Fetal DNA (Hücre Dışı Fetal DNA) |
| CMA | Chromosomal Microarray Analysis (Kromozomal Mikrodizin Analizi) |
| CS | Cesarean Section (Sezaryen) |
| CVS | Chorionic Villus Sampling (Koryon Villus Örneklemesi) |
| CP | Complexity Parameter (Karmaşıklık Parametresi) |
| CNV | Copy Number Variation (Kopya Sayısı Değişikliği) |
| DNN | Deep Neural Network (Çok Katmanlı Sinir Ağları) |
| DNA | Deoksiribonükleik Asit |
| EMQN | European Molecular Genetics Quality Network (Avrupa Moleküler Genetik Kalite Ağı) |
| HBYS | Hastane Bilgi Yönetim Sistemi |

| | |
|---|---|
| hCG | Human Chorionic Gonadotropin (İnsan Koryonik Gonadotropini) |
| ICP | Intracranial Pressure (İntrakraniyal-Kafa İçi Basıncı) |
| IUGR | Intrauterine Growth Restriction (İntrauterin Büyüme Geriliği) |
| Kb | Kilobaz |
| KKH | Congenital Heart Disease (Konjenital Kalp Hastalıkları) |
| KNN | K-Nearest Neighbors (K-En Yakın Komşular) |
| ML-MÖ | Machine Learning (Makine Öğrenmesi) |
| mAP | Mean Average Precision (Ortalama Kesinlik) |
| NA | Not Available/Not Applicable (Mevcut Değil/Uygulanamaz) |
| NF | Nuchal Fold (Ense Katlantısı) |
| NIPT | Non-İnvaziv Prenatal Testler |
| NT | Nuchal Translucency (Fetal Ense Kalınlığı) |
| PAPP-A | Pregnancy Associated Plasma Protein A (Gebelik ile İlişkili Plazma Proteini-A) |
| QF-PCR | Quantitative Fluorescence Polymerase Chain Reaction (Kantitatif Floresan Polimeraz Zincir Reaksiyonu) |
| ROC | Receiver Operating Characteristic (Alıcı İşlem Karakteristiği) |
| ROH/AOH | Regions of Homozygosity/Absence of Heterozygosity (Homozigotluk Bölgeleri/Heterozigotluk Yokluğu) |
| SCA | Sex Chromosome Aneuploidy (Cinsiyet Kromozomu Anöploidisi) |
| SNP | Single Nucleotide Polymorphism (Tek Nükleotid Polimorfizmi) |
| STR | Short Tandem Repeat (Kısa Tekrar Dizileri) |
| SS | Standart Sapma |
| SVM | Support Vector Machine (Destek Vektör Makinesi) |
| UPD | Uniparental Disomy (Tek Ebeveynli Dizomi) |
| USG | Ultrasonografi |
| VSD | Ventriküler Septal Defekt |
| WES | Whole Exome Sequencing (Tam Ekzom Dizileme) |

# ŞEKİL LİSTESİ


**Şekil 2.1** Prenatal tanı yöntemlerinin invaziv ve non-invaziv testler olarak sınıflandırılması . . . . . 7
**Şekil 2.2** QF-PCR analizlerinde kullanılan invaziv prenatal tanı yöntemlerinden elde edilen numune türleri . . . . . 10
**Şekil 3.1** QF-PCR kitinde bulunan 26 markır ve boy standart detayları . . . 14
**Şekil 3.2** Prenatal tanıda QF-PCR test sürecinin numune alımından raporlamaya kadar olan analiz aşamaları . . . . . 16
**Şekil 3.3** Kromozomal anomalilerin sınıflandırılması ve klinik sonuçları: trizomi, triploidi ve monozomi sendromları . . . . . 18
**Şekil 3.4** QF-PCR testi ile Trizomi 13 profil analizi . . . . . 18
**Şekil 3.5** QF-PCR testi ile Trizomi 18 profil analizi . . . . . 19
**Şekil 3.6** QF-PCR testi ile Trizomi 21 profil analizi . . . . . 20
**Şekil 3.7** QF-PCR testi ile Turner Sendromu profil analizi . . . . . 21
**Şekil 3.8** QF-PCR testi ile Klinefelter Sendromu profil analizi . . . . . 22
**Şekil 3.9** QF-PCR testi ile Triple X profil analizi . . . . . 23
**Şekil 3.10** QF-PCR testi ile Triploidi profil analizi . . . . . 24
**Şekil 5.1** Veri setinde eksik değerlerin tespiti, değer aralıklarının kontrolü ve tutarsızlıkların kontrolü için kullanılan R kodu . . . . . 37
**Şekil 5.2** Yaş dağılımını görselleştirmek için kullanılan R kodu . . . . . 47
**Şekil 5.3** Yaş ve QF-PCR sonuçlarına göre dağılımı görselleştirmek için kullanılan R kodu . . . . . 48
**Şekil 5.4** QF-PCR veri setinin makine öğrenmesi için %70 eğitim ve %30 test verisi olarak bölünmesi . . . . . 50
**Şekil 5.5** QF-PCR verilerinin sınıflandırılmasında kullanılan farklı modellerin ROC eğrileri ve Random Forest modeli için özellik önem analizi kodlaması . . . . . 51
**Şekil 7.1** Yaş dağılımını gösteren yoğunluk grafiği . . . . . 58
**Şekil 7.2** Yaşa göre QF-PCR sonuçlarının yoğunluk dağılımı . . . . . 60
**Şekil 7.3** Ultrason kategorilerine göre normal ve anomali QF-PCR sonuçlarının dağılım analizi . . . . . 66
**Şekil 7.4** Eğitim ve test veri setlerinin QF-PCR sonuçlarına göre dağılımı . 67

**Şekil 7.5** Ultrason verileri dahil edildiğinde makine öğrenmesi modellerinin ROC analizi karşılaştırması . . . . . 70
**Şekil 7.6** Ultrason verileri hariç tutulduğunda makine öğrenmesi modellerinin ROC analizi karşılaştırması . . . . . 71
**Şekil 7.7** Random Forest modelinde değişkenlerin Doğruluk Azalması ve Gini Saflık İndeksine göre önem dereceleri . . . . . 72
**Şekil 7.8** QF-PCR sonuçlarının tahmininde kullanılan Karar Ağacı modeli . 74
**Şekil 7.9** Prenatal tarama test sonuçlarının risk kategorilerine göre terminasyon ve gebelik sonucu tahmini üzerindeki etkisinin analizi 77
**Şekil 7.10** Prenatal tanıda makine öğrenmesi entegrasyonunun sistematik iş akış şeması . . . . . 79

# TABLO LİSTESİ


**Tablo 3.1** QF-PCR kitinde bulunan 26 markırın kromozom ve boy standart detayları . . . . . 15
**Tablo 3.2** Array-CGH ve CMA karşılaştırması . . . . . 27
**Tablo 5.1** Çalışmada kullanılacak parametreler ve açıklamaları . . . . . 35
**Tablo 5.2** Kompleks kromozamal varyasyonların kategorize edilmesi . . . . 44
**Tablo 5.3** USG sınıflandırma tablosu . . . . . 45
**Tablo 7.1** QF-PCR sonuçlarının sayıları, yaş ortalaması ve standart sapma . 60
**Tablo 7.2** Prenatal QF-PCR test sonuçlarının gebelik sonlanma şekillerine göre dağılımı . . . . . 62
**Tablo 7.3** Kromozomal anomalilerin QF-PCR sonuçlarına göre ebeveynsel kalıtım dağılımı . . . . . 62
**Tablo 7.4** Prenatal tarama programına katılan gebelerin demografik ve klinik özelliklerinin tanımlayıcı istatistikleri . . . . . 63
**Tablo 7.5** Prenatal tarama testi kullanımının yaş gruplarına göre dağılım analizi . . . . . 63
**Tablo 7.6** Maternal yaş gruplarına göre gebelik sonlanma şekillerinin dağılımı 64
**Tablo 7.7** Prenatal kromozomal anomali testlerinin (QF-PCR, Karyotipleme ve Array) tanısal uyum ve korelasyon analizi . . . . . 64
**Tablo 7.8** Prenatal ultrasonografi bulgularının normal ve anormal olarak yüzdesel dağılımı . . . . . 65
**Tablo 7.9** QF-PCR sonuçlarının sınıflandırılmasında %70 eğitim-%30 test veri seti ayrımı ile uygulanan makine öğrenmesi algoritmalarının performans metrikleri . . . . . 68

# ÖZET

## Prenatal Tanıda Trizomi Risk Tahmini için QF-PCR Sonuçları, Demografik, Klinik ve Genetik Verilerin Entegrasyonunun Makine Öğrenmesi Tabanlı Analizi

Kübra TÜREGÜN

Biyoteknoloji Anabilim Dalı
Biyoteknoloji Programı
Yüksek Lisans Tezi

Danışman: Doç. Dr. Alper YILMAZ

Prenatal tanı süreci, fetüste olası genetik ve kromozomal anomalileri gebelik sırasında tespit etmeyi hedefleyen kritik bir süreçtir. Özellikle Trizomi 21 (Down sendromu), Trizomi 18 (Edwards sendromu) ve Trizomi 13 (Patau sendromu) gibi yaygın kromozomal bozuklukların erken teşhisi, gebelik yönetimi ve fetal sağlık açısından büyük önem taşır. Bu çalışmada, QF-PCR (Kantitatif Floresan Polimeraz Zincir Reaksiyonu) yöntemi ile elde edilen sonuçlar, hastaların demografik (yaş, sigara kullanımı, hamilelik geçmişi), klinik (ultrasonografi bulguları, tarama test sonuçları) ve genetik verileriyle birlikte değerlendirildi. Bu veriler makine öğrenmesi yöntemleri ile entegre edilerek trizomi riskinin erken ve doğru tahmin edilmesi amaçlandı. Toplanan veriler, Lojistik Regresyon, Random Forest, Karar Ağacı ve K-Nearest Neighbors (KNN) gibi makine öğrenmesi algoritmaları kullanılarak analiz edildi. Ve modellerin performansları AUC, F1 skoru ve doğruluk gibi metriklerle ölçüldü.

Random Forest algoritması, AUC değeri 0.99 ve F1 skoru 0.87 ile en yüksek doğruluk oranına ulaşarak trizomi riskini tahmin etmede diğer algoritmalar arasında en iyi performansı gösterdi. Lojistik Regresyon algoritması güçlü bir doğruluk oranı (AUC: 0.86) sunarken, KNN algoritması beklenenin altında bir performans gösterdi (AUC: 0.52). Bu sonuçlar, yapay zeka ve makine öğrenmesi tabanlı yaklaşımların, prenatal dönemde yüksek doğruluk oranıyla trizomi riskini tahmin etmekte ve gereksiz invaziv testlerin önüne geçmede etkin bir araç olabileceğini göstermektedir.

Çalışmanın bulguları, prenatal tanıda trizomi risk tahminlerinin doğruluğunu artırmanın yanı sıra, bu tahminlerin klinik uygulamalara entegrasyonu için değerli içgörüler sunmaktadır. Bu modelleme, özellikle yüksek riskli gebeliklerde hızlı ve bilinçli kararlar alınmasını destekleyerek hasta yönetim sürecini optimize etmektedir. Makine öğrenmesi algoritmalarının prenatal tanı süreçlerine entegrasyonu, hem maddi hem de manevi açıdan sağlık sistemine katkı sağlamakta; genetik danışmanlık süreçleri, hasta yönetimi ve gelecekteki prenatal tanı yöntemlerine dair önemli ilerlemeler sunmaktadır.

**Anahtar Kelimeler:** Prenatal tanı, trizomi, QF-PCR, makine öğrenmesi, gebelik yönetimi.

# ABSTRACT


## Machine Learning Analysis of Demographic, Clinical and Genetic Data Integration with QF-PCR Results for Prenatal Diagnosis Trisomy Risk Prediction

Kübra TÜREGÜN

Department of Biotechnology
Master of Science Thesis

Supervisor: Assoc. Prof. Alper YILMAZ

Prenatal diagnosis is a critical process aimed at detecting potential genetic and chromosomal anomalies in the fetus during pregnancy. Early detection of common chromosomal disorders, particularly Trisomy 21 (Down syndrome), Trisomy 18 (Edwards syndrome) and Trisomy 13 (Patau syndrome) is crucial for pregnancy management and fetal health. In this study, results obtained through QF-PCR (Quantitative Fluorescent Polymerase Chain Reaction) were evaluated in conjunction with patients' demographic (age, smoking status and pregnancy history), clinical (ultrasonography findings and screening test results) and genetic data. These data were integrated with machine learning methods to achieve early and accurate prediction of trisomy risk. The collected data were analyzed using machine learning algorithms such as Logistic Regression, Random Forest, Decision Tree and K-Nearest Neighbors (KNN) and the models' performances were measured using metrics such as AUC, F1 score and accuracy.

The Random Forest algorithm demonstrated superior performance among other algorithms in predicting trisomy risk, achieving the highest accuracy rate with an AUC value of 0.99 and F1 score of 0.87. While the Logistic Regression algorithm provided a robust accuracy rate (AUC: 0.86), the KNN algorithm showed lower than expected performance (AUC: 0.52). These results indicate that artificial intelligence and machine learning-based approaches can serve as effective tools in predicting trisomy risk with high accuracy during the prenatal period and preventing unnecessary invasive tests.

The study's findings not only enhance the accuracy of trisomy risk predictions in prenatal diagnosis but also provide valuable insights into the integration of these predictions into clinical practice. This modeling optimizes patient management processes by supporting rapid and informed decision-making, particularly in high-risk pregnancies. The integration of machine learning algorithms into prenatal diagnostic processes contributes both financially and ethically to the healthcare system, offering significant advancements in genetic counseling processes, patient management and future prenatal diagnostic methods.

**Keywords:** Prenatal diagnosis, trisomy, QF-PCR, machine learning, pregnancy management.

# 1
# GİRİŞ

## 1.1 Literatür Özeti

Prenatal tanı, gebelik sürecinde fetüsteki genetik veya kromozomal anomalileri tespit etmeyi amaçlayan kritik bir süreçtir. Bu süreç, özellikle Trizomi 21 (Down sendromu), Trizomi 18 (Edwards sendromu) ve Trizomi 13 (Patau sendromu) gibi yaygın kromozomal bozuklukların erken aşamada belirlenmesini hedeflemektedir [1, 2]. Bu bağlamda, Kantitatif Floresan Polimeraz Zincir Reaksiyonu (QF-PCR) tekniği, prenatal tanıda devrim niteliğinde bir ilerleme sağlamıştır. QF-PCR, hızlı ve etkili bir yöntem olarak öne çıkmakta, fetal DNA üzerinde analiz yaparak kromozomal anomalileri tespit etmektedir [3, 4].

QF-PCR'nin en önemli avantajlarından biri, sonuçları geleneksel yöntemlere kıyasla çok daha kısa sürede, genellikle bir veya iki gün içinde sunabilmesidir. Bu hız, özellikle yüksek riskli gebeliklerde hızlı karar almayı kolaylaştırmaktadır [3, 5]. Ayrıca, QF-PCR'nin maliyet etkin olması ve yüksek doğruluk oranları sunması, bu tekniğin prenatal tanıda yaygınlaşmasının başlıca nedenlerindendir. Büyük ölçekli kliniklerde sıklıkla kullanılan bu test, yaygın kromozomal anomalilerin yanı sıra bazı nadir mozaisizm de tespit edebilme yeteneğine sahiptir [4]. Bununla birlikte, yapısal kromozomal anomalileri tespit etmede sınırlı kalabilmektedir, bu nedenle bazen karyotip analizi veya array-CGH gibi daha kapsamlı testlerle birleştirilmektedir [6].

Son yıllarda, makine öğrenmesi ve yapay zeka algoritmalarının prenatal tanıda kullanılması, QF-PCR gibi testlerden elde edilen verilerin daha etkin bir şekilde analiz edilmesini sağlamıştır. Bu yenilikçi yaklaşımlar, trizomi riskinin doğru bir şekilde tahmin edilmesine ve gereksiz invaziv testlerin önüne geçilmesine yardımcı olmaktadır. Örneğin, QF-PCR sonuçları ile demografik ve klinik verilerin makine öğrenmesi algoritmalarıyla entegrasyonu, kromozomal anomalilerin erken teşhisini ve gebelik yönetimini daha etkin hale getirmektedir [1, 2]. Özellikle yüksek riskli gebeliklerde, bu yapay zeka tabanlı modeller hızlı müdahalelere olanak tanıyarak klinik karar süreçlerini optimize etmekte ve tanı sürecini hızlandırmaktadır [7].

Son yıllarda, makine öğrenmesi ve yapay zeka algoritmalarının prenatal tanıda kullanımı, QF-PCR gibi testlerden elde edilen verilerin daha etkin analizine olanak sağlamaktadır. Bu yenilikçi yaklaşımlar, trizomi riskinin daha doğru tahmin edilmesini ve gereksiz invaziv testlerin önlenmesini mümkün kılmaktadır. Ancak mevcut literatürde, QF-PCR sonuçlarının klinik ve demografik verilerle bütünleştirilerek makine öğrenmesi tabanlı karar destek sistemlerine dönüştürüldüğü kapsamlı çalışmalar sınırlıdır.

Bu araştırma, literatürdeki bu boşluğu doldurmak amacıyla, QF-PCR analizlerini demografik ve klinik verilerle entegre eden yenilikçi bir yaklaşım sunmaktadır. Önerilen metodoloji, prenatal tanı sürecinde invaziv testlerin gerekliliğini azaltarak sağlık sistemine hem zaman hem de maliyet açısından önemli katkılar sağlama potansiyeli taşımaktadır. Ayrıca bu yaklaşım, QF-PCR'nin tanısal değerini artırırken, ailelere gebelik sürecinde daha bilinçli karar alma olanağı sunmakta ve fetüs sağlığının korunmasında kritik bir rol oynamaktadır.

## 1.2 Tezin Amacı

Prenatal tanıda trizomi riski, QF-PCR sonuçları ile yaş, tarama testleri, ultrasonografi, sigara kullanımı, hamilelik geçmişi ve abort sayısı gibi değişkenler göz önünde bulundurularak değerlendirilmekte ve bu faktörler hastanın terminasyon kararını etkileyebilmektedir. Terminasyon gibi kritik kararlar için risk değerlendirmelerinin hızlı, etkili ve çok yönlü bir şekilde yapılması gerekmektedir.

Bu çalışma, prenatal tanıda karar alma süreçlerini iyileştirmek amacıyla QF-PCR sonuçlarını demografik ve klinik verilerle entegre ederek makine öğrenmesi yaklaşımı geliştirmeyi hedeflemektedir. Bu kapsamda Lojistik Regresyon, Random Forest ve Karar Ağacı algoritmaları kullanılarak erken teşhis ve doğru risk değerlendirmesi sağlanması amaçlanmaktadır. Çalışma, yalnızca trizomi risk tahminlerini geliştirmeyi değil, aynı zamanda bu tahminlerin klinik uygulamalara nasıl entegre edilebileceğini ortaya koymayı da hedeflemektedir. Geliştirilen metodoloji ile invaziv testlerin gerekliliğinin azaltılması ve böylece sağlık sistemine hem ekonomik hem de operasyonel katkı sağlanması öngörülmektedir. Araştırmanın sonuçlarının, genetik danışmanlık süreçlerinin iyileştirilmesine, hasta yönetiminin geliştirilmesine ve prenatal tanı yöntemlerinin geleceğine önemli katkılar sunması beklenmektedir.

## 1.3 Hipotez

Klinik (ultrasonografi bulguları, tarama test sonuçları), demografik (yaş, hamilelik geçmişi, abort sayısı, sigara kullanımı) ve genetik (QF-PCR analizleri) verilerin makine öğrenmesi algoritmaları ile entegre edilmesi, prenatal tanıda trizomi (13, 18 ve 21) riskinin erken ve yüksek doğrulukla tahmin edilmesini mümkün kılar. Bu tahminlerin, mevcut tanı yöntemlerinin doğruluğunu artırarak, gebelik yönetiminde erken müdahaleyi kolaylaştıracağı ve terminasyon kararları gibi kritik klinik süreçlerde daha bilinçli ve etkili kararlar alınmasını sağlayacağı öngörülmektedir. Ayrıca, bu yaklaşım gereksiz invaziv testlerin önüne geçerek sağlık sistemine hem maddi hem de manevi açıdan katkı sunabilir.

Bu çalışmada kullanılan makine öğrenmesi algoritmaları, yalnızca yüksek doğruluk oranları sunmakla kalmayacak, aynı zamanda modelin klinik uygulamalara entegrasyonuyla, genetik danışmanlık ve prenatal tanı süreçlerinde paradigma değişikliğine yol açacaktır. Çalışmanın bulgularının, genetik ve klinik verilerin birlikte kullanıldığı disiplinler arası yaklaşımların önemini vurgulaması ve bu alandaki gelecekteki araştırmalara yol göstermesi beklenmektedir.

# 2
# PRENATAL TANI VE PRENATAL TANI TESTLERİ

## 2.1 Prenatal Tanı

Prenatal tanı, gebelik sırasında fetüsün genetik ve yapısal özelliklerini değerlendirerek kalıtsal hastalıklar ve kromozom bozukluklarını erken dönemde tespit etmeyi amaçlar. Bu tanı süreci, fetüsün sağlık durumunu anlamak adına hem invaziv (örn. amniyosentez, koryon villus biyopsisi) hem de non-invaziv (örn. hücre dışı fetal DNA [cffDNA] analizi) testleri içerir. İnvaziv yöntemler daha doğrudan veri sağlarken düşük riski gibi komplikasyonlara yol açabilir. Buna karşın, non-invaziv prenatal testler (NIPT) anne kanında dolaşan cffDNA'yı analiz ederek trizomi gibi kromozom bozukluklarını tespit edebilme kapasitesine sahiptir ve düşük riski taşımadığı için sıklıkla tercih edilir [8–10].

Prenatal tanı uygulamaları, özellikle Down sendromu (Trizomi 21), Edward sendromu (Trizomi 18) ve Patau sendromu (Trizomi 13) gibi kromozom anomalilerini yüksek güvenilirlikte tarama imkanı sunar. Bu testler, özellikle ileri anne yaşı, ailede genetik bozukluk öyküsü veya ultrason bulgularında anormallikler gibi yüksek riskli gebeliklerde büyük önem taşır. Biyokimyasal taraması ve ultrason gibi non-invaziv yöntemler yanı sıra, amniyosentez ve koryon villus biyopsisi gibi invaziv testlerle doğrudan fetal DNA'ya erişim sağlanabilir. Bu tanılar, fetüsün gelişimsel sağlığına dair kapsamlı bilgi sunarak doğum öncesinde yapılacak müdahalelerin planlanmasına yardımcı olur [3, 4].

Son yıllarda NIPT'in hızlı gelişimi, yalnızca anöploidi taramasında değil, aynı zamanda tek gen bozuklukları gibi karmaşık genetik hastalıkların belirlenmesinde de etkin hale gelmiştir. Ancak non-invaziv yöntemlerin güvenilirliği teyit edilmek istendiğinde invaziv testler hala gereklilik arz edebilir [11, 12]. Örneğin, Kantitatif Floresan Polimeraz Zincir Reaksiyonu (QF-PCR) gibi hızlı prenatal testlerdeki ilerlemeler, sonuç alma süresini önemli ölçüde kısaltarak gebelik sırasındaki karar verme sürecini iyileştirmiştir [4, 13].

Prenatal tanı, ayrıca bazı etik ve sosyal zorlukları da beraberinde getirir. Örneğin, NIPT testleri nadir durumlarda fetüste değil de anne adayında kanser tespit edilmesi gibi belirsizlikler yaratabilir. Elde edilen sonuçların ebeveynlere sağlanması ve gebelik yönetiminde bilinçli kararların alınmasına katkı sağlarken, genetik danışmanlık süreci de ebeveynlerin testlerin fayda ve sınırlamaları hakkında bilgilendirilmesi açısından önem taşır [14, 15].

Gelişen teknolojiler sayesinde prenatal tanı, sadece karyotipleme gibi geleneksel yöntemlerle sınırlı kalmamakta, array tabanlı Karşılaştırmalı Genomik Hibridizasyon (array-CGH) ve Tam Ekzom Dizileme (WES) gibi ileri yöntemlerle kromozomal anormallikleri çok daha ince bir çözünürlükle tespit etmektedir [5, 16]. Bu yeni nesil teknikler, genetik yapının daha ayrıntılı anlaşılmasını sağlamakta, teşhis doğruluğunu artırmakta ve özel bakım gereksinimleri için ailelerin hazırlık yapmasına yardımcı olmaktadır [6].

Prenatal tanı, gebelik sırasında hem invaziv hem de non-invaziv yöntemlerle fetüsün genetik sağlığını değerlendirme imkanı sunar. Yüksek riskli gebeliklerde genetik danışmanlık ve erken tanı, bilinçli karar verme süreçlerini destekler ve doğum sonrası özel bakım gereksinimlerine hazırlanma sürecini kolaylaştırır. Gelişen genetik teknolojiler, tanı doğruluğunu artırarak prenatal tanının değerini güçlendirmektedir.

## 2.2 Gebelik Dönemleri ve Özellikleri

Gebelik süresi genellikle üç aylık periyotlara bölünür ve bu dönemler trimester olarak adlandırılır. Toplam gebelik süresi yaklaşık 40 hafta olup, üç trimesterden oluşur. Her trimesterin kendine özgü fizyolojik değişiklikleri ve fetüs gelişim aşamaları vardır.

### 2.2.1 Birinci Trimester (1-13. Hafta)

Birinci trimester, gebeliğin başlangıcından 13. haftaya kadar olan süreyi kapsar. Bu dönem, fetüsün organ sistemlerinin oluşmaya başladığı ve kritik bir gelişim aşamasıdır.

**Fizyolojik değişiklikler:** Anne adayında sabah bulantıları, yorgunluk ve hormonal değişiklikler sık görülür. Beta-hCG hormonu bu dönemde zirve yapar.

**Fetüs gelişimi:** İlk 8 hafta boyunca embriyo olarak adlandırılan fetüs, bu dönemin sonunda temel organ ve doku yapılarını tamamlar. Kalp atışları genellikle 6. haftadan itibaren ultrasonla tespit edilebilir. 12. haftada fetüs yaklaşık 7-8 cm uzunluğa ulaşır

ve 14 gram ağırlığındadır. Düşük riskinin en yüksek olduğu dönemdir [17, 18].

### 2.2.2 İkinci Trimester (14-26. Hafta)

İkinci trimester, genellikle gebeliğin en rahat dönemi olarak kabul edilir. Bu dönemde anne adayının enerji seviyeleri artar ve fetal hareketler hissedilmeye başlanır.

**Fizyolojik değişiklikler:** Karın büyümeye başlar ve rahim pelvisin üzerine çıkar. Anne adayında sırt ağrıları, çatlaklar ve hafif ödem gibi belirtiler görülebilir. Gebelik semptomlarının genellikle hafiflediği dönemdir.

**Fetüs gelişimi:** Organlar işlevsel hale gelir; örneğin böbrekler idrar üretir ve sinir sistemi gelişir. Fetüs, annenin sesini duyabilir ve ışığa tepki verebilir. 24. haftaya gelindiğinde fetüs yaklaşık 600 gram ağırlığındadır, hayatta kalma şansı artar ve cinsiyet tayinin yapıldığı dönemdir [19, 20].

### 2.2.3 Üçüncü Trimester (27-40. Hafta)

Üçüncü trimester, doğuma hazırlık dönemidir. Bu dönemde fetüs hızla kilo alır ve gelişimini tamamlar.

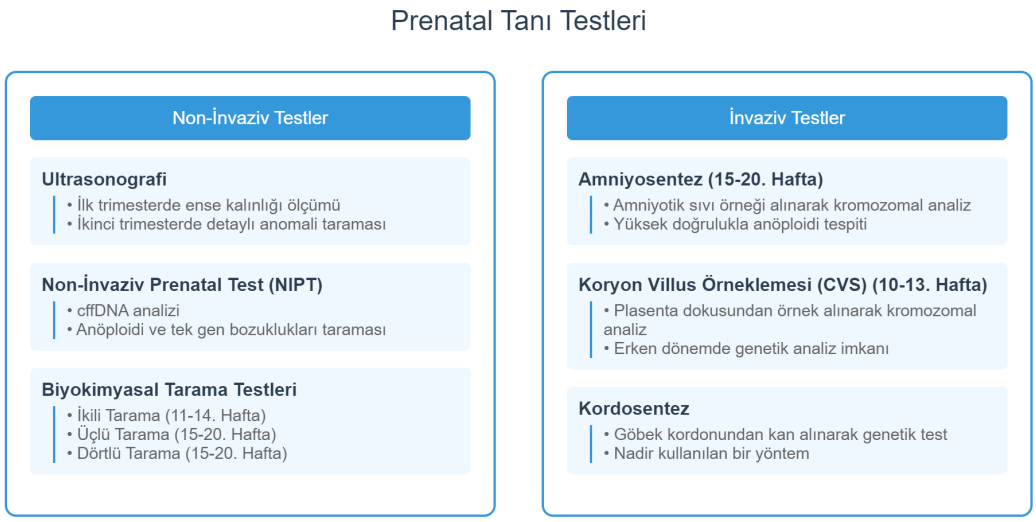
**Fizyolojik değişiklikler:** Anne adayında sık idrara çıkma, nefes darlığı ve rahim kasılmaları (Braxton Hicks) görülebilir. Vücut, doğum için hazırlık yapar; pelvik kemikler esnemeye başlar.

**Fetüs gelişimi:** Beyin hızla büyür ve bağışıklık sistemi gelişir. Akciğerler doğum sonrası nefes almaya hazırlanır. Doğumdan hemen önce fetüs genellikle 2.5-4 kg arasında ağırlığa ve yaklaşık 50 cm uzunluğa ulaşır [18–20].

Trimesterler, gebelikte hem anne hem de fetüs için önemli dönüm noktalarını ifade eder. Her dönem farklı fizyolojik ve biyolojik değişimleri içerir ve bu süreçlerde yapılan tarama testleri ile fetüsün sağlığı yakından izlenir.

## 2.3 Prenatal Tanı Testleri

Prenatal tanı testleri, gebelik süresince fetüsün sağlık durumu hakkında bilgi sunan farklı tarama ve tanı testlerini içerir. Bu testler, genetik hastalıkların erken teşhisine olanak tanıyarak, ebeveynlere bilinçli kararlar alma fırsatı sunar ve gebeliğin uygun şekilde yönetilmesine katkıda bulunur [8, 21]. Prenatal tanı testlerinin non-invaziv ve invaziv testler olarak sınıflandırılması Şekil 2.1'de gösterilmektedir.




**Şekil 2.1** Prenatal tanı yöntemlerinin invaziv ve non-invaziv testler olarak sınıflandırılması

Prenatal tanı testleri, tarama ve tanı amaçlı uygulanan çeşitli testleri kapsar. İlk olarak ultrasonografi, fetüsün fiziksel yapısının incelenmesi ve bazı yapısal anomalilerin erken teşhisi için önemli bir rol oynar. İlk trimesterde yapılan detaylı ultrasonografi, özellikle ense kalınlığı (nt, nuchal translucency) ölçümü ile Down sendromu gibi anormallikler hakkında ön bilgi verir. İkinci trimesterde ise detaylı anomali taraması gerçekleştirilir [8, 22].

Biyokimyasal testler de prenatal tanıda sıkça kullanılır. İkili, üçlü ve dörtlü tarama testleri olarak bilinen bu testler, anne kanında bulunan belirli protein ve hormonların düzeylerini analiz eder. İkili tarama testi, 11-14. haftalarda PAPP-A ve serbest $\beta$-hCG değerlerine dayanarak fetüsün Down sendromu gibi genetik anormallik riskini belirler. Üçlü ve dörtlü tarama testleri ise ikinci trimesterde uygulanır ve fetüsün anöploidi riskini değerlendirmeye katkı sağlar [12, 23, 24].

Non-invaziv testlerin en önemli avantajı, anne kanından cffDNA örneği alarak fetal kromozom anormalliklerini saptayabilmesi ve invaziv testlere duyulan ihtiyacı azaltmasıdır [9, 15]. İnvaziv yöntemler ise amniyosentez, kordosentez ve koryon villus örneklemesi (CVS) gibi testleri içerir. Bu yöntemler daha yüksek doğruluk sağlar ancak düşük riski gibi komplikasyonlar oluşturabilir. Amniyosentez, genellikle 15-20. haftalar arasında gerçekleştirilir ve amniyotik sıvı örneği alarak kromozomal anormallikleri tanımlamada kullanılır. CVS ise daha erken, 10-13. haftalarda yapılabilir ve plasentadan alınan hücrelerle genetik analiz sağlar. Kordosentez, ikinci veya üçüncü trimesterde kordon kanından örnek alarak yapılan bir testtir ve nadiren kullanılır [12, 25, 26].

### 2.3.1 Non-İnvaziv Testler

#### 2.3.1.1 Ultrasonografi Taramaları

Ultrasonografi, prenatal tanının temel taşlarından biridir ve gebelik boyunca farklı dönemlerde gerçekleştirilir. İlk trimesterde yapılan ultrasonografi, fetüsün ense kalınlığı ölçümü ile Down sendromu gibi kromozom anomalilerini erken evrede tespit etme potansiyeline sahiptir [11, 25]. İkinci trimesterde detaylı bir anomali taraması yapılır; bu tarama, kalp anomalileri, omurga defektleri ve ekstremitelerdeki yapısal sorunlar gibi pek çok fiziksel anomaliyi ortaya çıkarabilir [8, 22]. Ultrasonografi taramaları, yüksek riskli gebeliklerde daha ileri değerlendirmeler için temel bir yol gösterici niteliğindedir.

#### 2.3.1.2 Non-İnvaziv Prenatal Test (NIPT)

NIPT, non-invaziv bir tarama yöntemi olup, anne kanında dolaşan cffDNA analizini içerir. Yüksek hassasiyet ve özgüllükle Trizomi 21 (Down sendromu), Trizomi 18 (Edward sendromu) ve Trizomi 13 (Patau sendromu) gibi yaygın kromozomal bozuklukları tespit edebilir [21, 26]. Özellikle yüksek riskli gebeliklerde tercih edilen bu test, pozitif sonuç verdiğinde doğrulayıcı invaziv testlerle desteklenmelidir. NIPT'in avantajı, düşük riski olmadan fetüs hakkında kapsamlı bilgi sağlamasıdır; bu nedenle dünya genelinde rutin bir tarama yöntemi olarak kabul edilmektedir [9, 22].

#### 2.3.1.3 Biyokimyasal Tarama Testleri

Prenatal tarama testlerinden biyokimyasal testler, fetal anomalilerin erken dönemde saptanmasında kullanılan non-invaziv yöntemlerdir. Bu testler, gebeliğin farklı trimesterlerinde uygulanmakta ve çeşitli biyokimyasal belirteçleri içermektedir. Her test, spesifik gebelik haftalarında optimal sonuçlar vermektedir. Bu testler, riskli gebelikleri belirlemeye yardımcı olur ve gerektiğinde invaziv tanı testlerine yönlendirme sağlar.

Biyokimyasal tarama testleri, maternal kan (anne kan) örneğinde bulunan belirli biyomarkırların ölçümü yoluyla fetüsün anöploidi riski hakkında bilgi sunar. Bu testler şu şekilde sınıflandırılabilir:

**İkili Tarama (11-14. Hafta):** Gebeliğin ilk trimesterinde (11-14. haftalar arasında) yapılan bir testtir. Bu test, anne kanında **PAPP-A** ve **serbest $\beta$-hCG** düzeylerini ölçerek Down sendromu gibi genetik anormallik risklerini değerlendirir [12, 27]. Ultrason parametrelerinden, **fetal ense kalınlığı (NT)** değerlendirilir. Bu test, özellikle Down sendromu (Trizomi 21) ve Edwards sendromu (Trizomi 18) riskini

hesaplamak için kullanılır ve tarama doğruluğu yüksektir. İkili tarama testinin avantajı, erken sonuç vermesi ve risk yönetiminde hızlı hareket imkanı sağlamasıdır [17, 28]. Fetüsün nöral tüp defekti gibi diğer anomaliler için bir bilgi sağlamasa da, yüksek duyarlılık ve özgüllük sunar [18, 29].

**Üçlü Tarama (15-20. Hafta):** İkinci trimesterde uygulanan bu test, AFP (alfa-fetoprotein), hCG ve estriol düzeylerine bakar. Bu değerler, Down sendromu ve nöral tüp defektleri gibi durumlar için risk hesaplamasında kullanılır [11, 25].

**Dörtlü Tarama (15-20. Hafta):** Üçlü tarama parametrelerine inhibin-A eklenir, bu da testin Down sendromu gibi bazı durumları belirlemedeki doğruluğunu artırır. Dörtlü tarama, daha ileri bir risk değerlendirmesi sağlamak amacıyla tercih edilir [15, 26].

### 2.3.2 İnvaziv Testler

İnvaziv prenatal testler, fetüsün genetik durumunu doğrudan değerlendirmek amacıyla fetal hücrelerin elde edilmesine dayanan tanısal prosedürlerdir. Bu testler arasında en yaygın kullanılan yöntemler amniyosentez ve koryonik villus örneklemesi (CVS)'dir. Bu yöntemlerle elde edilen hücreler, fetüste Down sendromu (trizomi 21), Edwards sendromu (trizomi 18) ve Patau sendromu (trizomi 13) gibi kromozomal anomalilerin kesin teşhisinde kullanılır [30, 31].

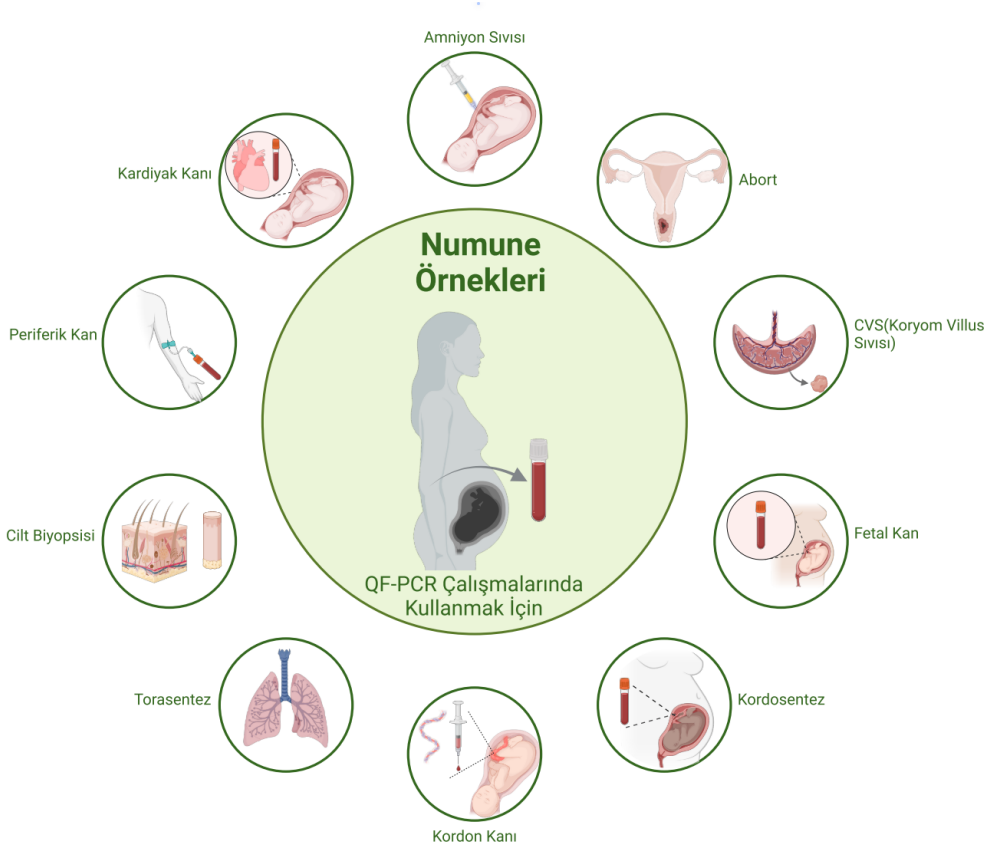
İnvaziv testler %99'a varan doğruluk oranlarına sahiptir, ancak bu prosedürler küçük de olsa komplikasyon riski taşır. Örneğin, amniyosentez ve CVS için düşük riski %0,3-0,5 arasında değişmektedir [30]. Bu nedenle, testler genellikle ultrason bulgularında anomali tespiti, tarama testlerinde yüksek risk ya da ileri anne yaşı gibi durumlarda önerilmektedir [31, 32].

Son yıllarda, non-invaziv prenatal testlerin (NIPT) gelişmesiyle invaziv testlerin kullanım oranı azalmış olsa da, NIPT'de saptanan anormalliklerin doğrulanması için invaziv yöntemler halen altın standart olma özelliğini korumaktadır [32].

#### 2.3.2.1 İnvaziv Genetik Tanıda Kullanılan Numune Örnekleri

İnvaziv genetik tanıda kullanılan numune örnekleri, invaziv prenatal tanı testleri için fetüsten doğrudan alınan hücre örneklerinin genetik analizini içerir ve yüksek doğruluk oranlarına sahiptir. İnvaziv genetik tanıda kullanılan numune örnekleri çeşitlilik göstermektedir. Bu testler için kullanılanlar Şekil 2.2'de gösterilen numuneler tipleridir.

**Amniyosentez:** Gebeliğin 15-20. haftaları arasında yapılır ve amniyotik sıvıdan alınan örnekte fetal hücrelerin kromozomal analizi gerçekleştirilir. Amniyosentez, özellikle NIPT veya biyokimyasal testlerin yüksek risk gösterdiği durumlarda önerilir ve %99,5 doğruluk oranına sahiptir [8, 10, 33, 34]. Amniyosentez, genetik tanı için yüksek riskli gebeliklerde yaygın olarak kullanılan bir yöntemdir. Ultrason rehberliğinde amniyotik sıvıdan alınan örnekler, fetal karyotip ve genetik analizlerde altın standart olarak kabul edilir [30, 32].




**Şekil 2.2** QF-PCR analizlerinde kullanılan invaziv prenatal tanı yöntemlerinden elde edilen numune türleri, BioRender ile oluşturuldu.

**Koryon Villus Örneklemesi (CVS):** Gebeliğin 10-13. haftalarında yapılır ve plasenta dokusundan alınan örnekle fetüsün kromozom yapısı incelenir. CVS, daha erken uygulanabilmesi nedeniyle avantaj sağlar ancak düşük riskini de içerir [12, 22, 33, 34]. Koryonik villus örneklemesi (CVS) ise plasental dokudan örnek alınarak yapılır ve erken gebelikte uygulanması nedeniyle bazı durumlarda tercih sebebi olur [30, 32].

**Kordosentez:** Gebeliğin ikinci veya üçüncü trimesterinde, fetüsün göbek kordonundan alınan kan örneği ile yapılır ve daha az yaygın kullanılır. Kordosentez, genellikle fetüste kanla ilgili genetik durumları incelemek için tercih edilir [27, 33].

**Kordon kanı:** Doğum sırasında bebekten gelen göbek kordonu içindeki kandır. Kök hücre açısından zengin olup, genetik incelemeler ve tıbbi uygulamalarda kullanılır [35].

Kordon kanı ve kordosentez ile elde edilen kan benzer görünse de aralarında önemli farklar vardır. Kordon kanı, doğum sırasında göbek kordonundan alınan kandır ve fetüsün dolaşımından bağımsızdır. Buna karşılık, kordosentez sırasında alınan kan doğrudan fetüsün dolaşımından veya plasentadan elde edilir. Bu nedenle, kordosentez fetüsün kanını doğrudan analiz etmek için kullanılırken, kordon kanı daha çok kök hücre açısından zengin içeriği ve genetik incelemeler için tercih edilir.

**Fetal kan:** Doğrudan fetüsten alınan kandır. Genellikle kordosentez veya amniyosentez gibi invaziv işlemlerle elde edilir ve fetal genetik hastalıkların tanısı için kullanılır [33, 35].

**Abort materyali:** Düşük (abortus) sonrasında rahimden elde edilen embriyo veya fetüs kalıntılarıdır. Genetik anomali incelemelerinde kullanılabilir [33].

**Cilt biyopsisi:** Fetüsün veya bireyin cilt dokusundan alınan biyopsidir. Genetik veya diğer hastalıkların tanısında kullanılabilir.

**Periferik kan:** Genellikle bireyin vücudundan, genellikle koldan alınan kandır. Yetişkin veya çocuklardan alınan kan örnekleri için kullanılır, fetüsten alınmaz [35].

**Kardiyak kan:** Kalpten alınan kandır. Fetüsten veya yetişkinlerden alınabilir. Fetüsün kardiyak kanı, fetal kan kategorisine girebilir [35].

**Torasentez:** Göğüs boşluğundan (plevra boşluğu) sıvı alma işlemidir. Genellikle akciğer hastalıklarının tanısı için yapılan bir işlemdir ve prenatal tanıdan farklıdır [36].

# 3
# KROMOZOMAL ANOMALİLERİN TANISINDA QF-PCR VE DİĞER YÖNTEMLER

Kromozomal anomaliler, prenatal tanıda kritik bir araştırma alanı olup, gebelik sonuçlarını doğrudan etkileyen ve klinik süreçte önemli zorluklar yaratan genetik bozukluklardır. Bu anomaliler arasında en sık karşılaşılanlar, Trizomi 21 (Down sendromu), Trizomi 18 (Edwards sendromu) ve Trizomi 13 (Patau sendromu) gibi yaygın trizomilerdir. İlk dönemlerde, bu anomalilerin tespiti maternal yaş gibi basit risk faktörlerine dayandırılırken, günümüzde tarama ve tanı süreçlerinde önemli gelişmeler yaşanmaktadır. Birinci trimester taraması, maternal kan biyobelirteçlerinin (ör. serbest beta-hCG, PAPP-A) ve fetal ense kalınlığı (NT) gibi ultrasonografik parametrelerin analizini içeren bir yöntem olarak, yaygın trizomilerin yüksek doğrulukla tespit edilmesinde önemli bir araç haline gelmiştir [17]. Buna ek olarak, hücre dışı fetal DNA (cffDNA) analizi, minimal invaziv bir yöntem olarak, yüksek sensitivite ve düşük yanlış pozitif oranları sunarak prenatal tarama yaklaşımlarında altın standart olarak kabul edilmektedir [17, 37]. Ancak, cffDNA analizinin, nadir kromozomal anomaliler veya yapısal fetal kusurları yeterince kapsamaması gibi sınırlamaları bulunmaktadır, bu da daha ayrıntılı tanı yöntemlerinin kullanılmasını gerekli kılmaktadır [17, 37].

QF-PCR, prenatal tanıda trizomi 13, 18, 21 ve cinsiyet kromozomlarına bağlı anöploidilerin hızlı ve doğru bir şekilde tespit edilmesini sağlar. Polimorfik bölgelerin çoğaltılmasına dayalı bu yöntem, yalnızca birkaç saat içinde sonuç vererek gebelik yönetimi için kritik bir zaman kazancı sunar. Ayrıca, maternal hücre kontaminasyonu ve mozaisizm gibi sorunların üstesinden gelme yeteneği, QF-PCR'ı yüksek hacimli tanı uygulamaları için ideal bir araç haline getirmektedir [13]. Bu özellikleriyle QF-PCR, kromozomal anomalilerin erken ve kesin tanısında klinik değeri yüksek bir yöntem olarak öne çıkmaktadır ve prenatal genetik tanı süreçlerinde vazgeçilmez bir bileşen haline gelmiştir [13, 17].

## 3.1 Kromozomal Anöploidi Tanısında QF-PCR Metodolojisi

Kantitatif Floresan Polimeraz Zincir Reaksiyonu (QF-PCR), hedef kromozomlar üzerindeki kısa tekrar dizilerinin (STR) floresan işaretli primerler kullanılarak amplifikasyonuna dayanan, sayısal kromozom anomalilerinin tespiti için kullanılan bir yöntemdir. Trizomi 21 (Down Sendromu), Trizomi 18 (Edward Sendromu), Trizomi 13 (Patau Sendromu) gibi kromozomal anomalilerin yanı sıra Turner ve Klinefelter Sendromları gibi cinsiyet kromozomu anormalliklerinin teşhisi için de tasarlanmıştır.

QF-PCR, 13, 18, 21, X ve Y kromozomlarını etkileyen yaygın anöploidilerin tespitinde sağladığı hız ve doğruluk nedeniyle prenatal tanı alanında giderek daha fazla tercih edilen bir moleküler tanı aracıdır [3, 13]. Kültüre edilmiş fetal hücreler gerektiren ve sonuçların iki hafta kadar sürebildiği geleneksel karyotiplemenin aksine, QF-PCR, hedeflenen kromozomlar üzerindeki STR işaretleyicilerini amplifiye ederek bir ya da iki gün içinde sonuç verebilmektedir. Bu hızlı sonuç verme süresi, yüksek riskli gebeliklerde kritik kararlar almak için zamanında bilgiye ihtiyaç duyan ebeveynler için özellikle faydalıdır [4].

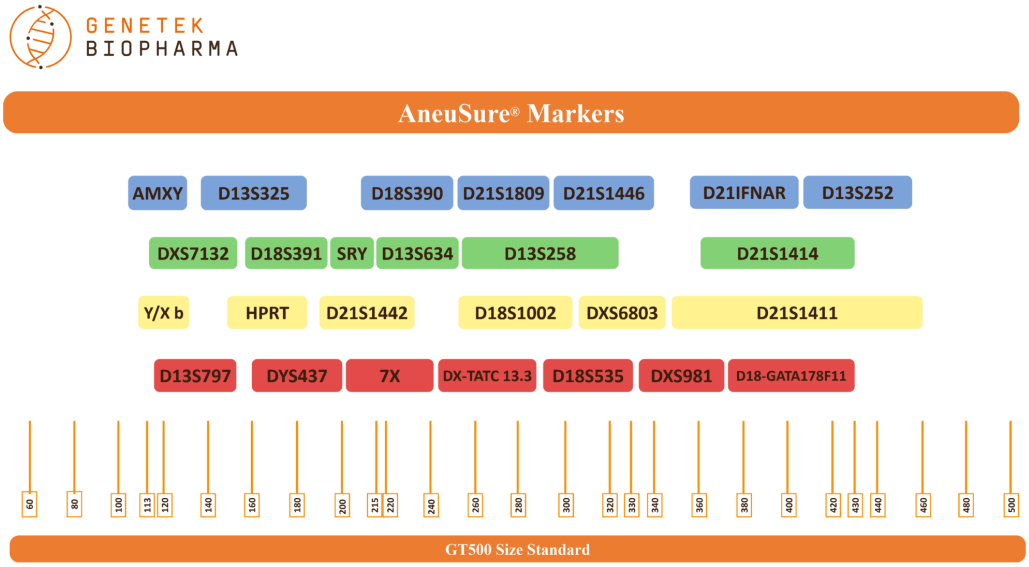
QF-PCR, kromozomal anomalilerin hızlı ve doğru tespitini sağlayan bir yöntemdir; ancak düşük düzeydeki mozaisizmi veya anne hücresi kontaminasyonunu tespit etme konusunda sınırlamalar gösterebilir. Bu durumlar, özellikle diğer tekniklerle birlikte değerlendirildiğinde tanısal doğruluğu etkileyebilir [13].

İşlem, spesifik kromozom bölgelerine bağlanan floresanla işaretlenmiş primerler kullanılarak gerçekleştirilir ve DNA fragmanları elektroforez ile analiz edilerek kromozom kopya sayısı belirlenir. Bu yöntem son derece hassas ve spesifiktir ve sonuçları genellikle tam karyotip analizleri ile uyumludur [3, 5]. Bununla birlikte, QF-PCR'nin translokasyonlar veya küçük duplikasyonlar gibi yapısal kromozomal anormallikleri tespit edememesi gibi bazı sınırlamaları vardır; bu tür anomaliler, Array-CGH gibi daha kapsamlı tekniklerle daha iyi tespit edilir [13, 16].

QF-PCR, sınırlamalarına rağmen, gebelik yönetiminde hızlı tanı gerektiğinde güvenilir ve maliyet etkin bir seçenek olarak kalmaya devam etmektedir [38]. Birçok durumda, en yaygın kromozom anöploidilerini tespit etmek için bağımsız bir test olarak kullanılır ve karyotipleme gibi daha iş gücü yoğun yöntemlerin kullanılma gereksinimini azaltır [5]. Bu nedenle, QF-PCR, prenatal tanıyı hem sağlık sağlayıcılar hem de bekleyen ebeveynler için daha erişilebilir ve verimli hale getirmektedir [4].

Çalışmamızda kullandığımız Genetek Biopharma AneuSure® QF-PCR kitinin

metodolojisi, kromozomal anöploidilerin tespitinde 13, 18, 21, X ve Y kromozomlarına özgü polimorfik mikrosatellit markırlarının analizine dayanmaktadır. STR lokuslarının selektif amplifikasyonuna dayanan bu metodoloji, kromozom-spesifik markırların kullanımını içermektedir. Şekil 3.1'de, AneuSure® QF-PCR kitinde bulunan 26 markır ve boy standart detayları, spesifik markırların konumsal dağılımını ve GT500 boy standardına göre baz çifti uzunluklarını göstermektedir. Markırlar dört farklı floresan boya ile işaretlenmiş (mavi, yeşil, sarı ve kırmızı paneller) olup, 90-450 baz çifti aralığında dağılım göstermektedir. Bu organize markır dizilimi, multipleks QF-PCR analizinde kromozomal anomalilerin hassas tespitine olanak sağlamaktadır [39].




**Şekil 3.1** AneuSure® QF-PCR kitinde bulunan 26 markır ve boy standart detayları [39].

Tablo 3.1'deki, her bir markırın kromozomal lokalizasyonunu, baz çifti cinsinden boyut aralığını ve kromozom üzerindeki spesifik konumunu göstermektedir. Bu markırlar, dört farklı floresan boya (mavi, yeşil, sarı ve kırmızı) ile işaretlenmiş olup, 13., 18., 21., X ve Y kromozomlarına özgü bölgeleri hedeflemektedir [39]. Sistem, markırları, boy standardını (size standardını) ve markırların sistematik uygulanmasını içermektedir [39]. Boy standardı, DNA fragmanlarının baz çifti uzunluklarını doğru bir şekilde belirlemek için kullanılan kalibrasyon aracıdır. Bu standart sayesinde, analiz edilen DNA parçalarının uzunlukları hassas bir şekilde ölçülebilmekte ve böylece kromozomal anomalilerin tespiti daha güvenilir hale gelmektedir. GT500 boy standardı, 35-500 baz çifti aralığında referans noktaları sağlayarak, incelenen DNA fragmanlarının boyutlarının kesin olarak belirlenmesine olanak tanır. STR markırlarının çoğaltılması, Avrupa Moleküler Genetik Kalite Ağı (EMQN) standartlarına uygun olarak floresan işaretli primerler ile gerçekleştirilir.

**Tablo 3.1** AneuSure® QF-PCR kitinde bulunan 26 markırın kromozom ve boy standart detayları [39].

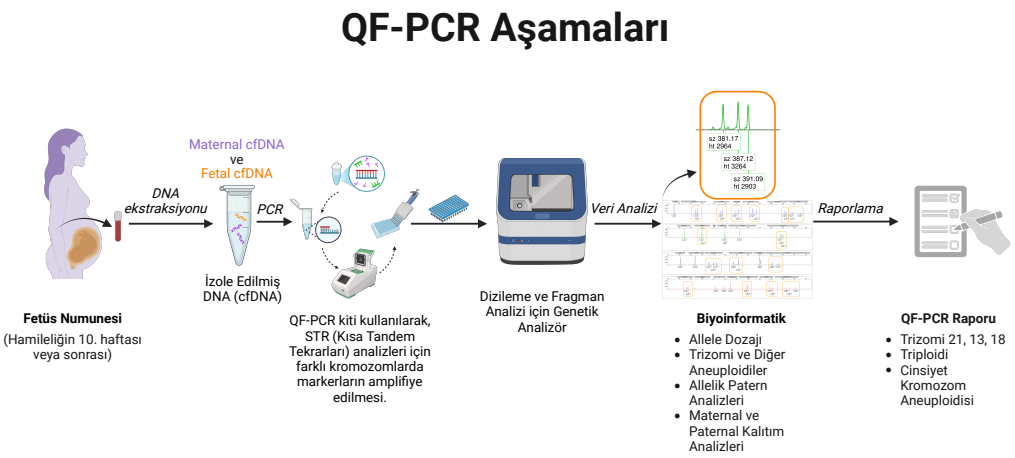
| No. | Markır | Kromozom | Boy Aralığı | Kr. Konumu |
|---|---|---|---|---|
| 1 | AMXY | X,Y | 102-120 | Xp22.2<br>Yp11.2 |
| 2 | D13S325 | 13 | 125-190 | 13q14.11 |
| 3 | D18S390 | 18 | 210-250 | 18q22.3 |
| 4 | D21S1809 | 21 | 260-280 | 21q22.2 |
| 5 | D21S1446 | 21 | 285-340 | 21q22.3 |
| 6 | D21IFNAR | 21 | 350-408 | 21q22.1 |
| 7 | D13S252 | 13 | 409-455 | 13q12.2 |
| 8 | DXS7132 | X | 115-150 | Xq12 |
| 9 | D18S391 | 18 | 160-200 | 18p11.31 |
| 10 | SRY | Y | 208-219 | Yp11.31 |
| 11 | D13S634 | 13 | 220-251 | 13q21.33 |
| 12 | D13S258 | 13 | 253-325 | 13q21.33 |
| 13 | D21S1414 | 21 | 350-430 | 21q21.1 |
| 14 | Y/X b | Y, X | 110-135 | Yp11.2<br>Xq21.31 |
| 15 | HPRT | X | 145-185 | Xq26.3 |
| 16 | D21S1442 | 21 | 188-233 | 21q21.3 |
| 17 | D18S1002 | 18 | 250-310 | 18q11.2 |
| 18 | DXS6803 | X | 315-344 | Xq21.31 |
| 19 | D21S1411 | 21 | 346-460 | 21q22.3 |
| 20 | D13S797 | 13 | 115-155 | 13q33.2 |
| 21 | DYS437 | Y | 160-200 | Yq11.21 |
| 22 | 7X | 7, X | 210-243 | 7q34<br>Xq13.3 |
| 23 | DX-TATC 13.3 | X | 245-271 | Xp21.2 |
| 24 | D18S535 | 18 | 275-325 | 18q12.3 |
| 25 | DXS981 | X | 330-365 | Xq13.1 |
| 26 | D18-GATA178F11 | 18 | 370-430 | 18p11.32 |

Metodolojik yaklaşımda, optimize edilmiş termal siklus cihazları kullanılarak spesifik DNA dizilerinin amplifikasyonu gerçekleştirilmektedir. Aşamaları Şekil 3.2'deki diyagramda gösterilmiştir; bu süreç hamileliğin 10. haftası veya sonrasında alınan fetüs numunesinden başlayarak DNA ekstraksiyonu ile devam eder. Bu süreç, primer-spesifik PCR protokollerinin standardizasyonu ile desteklenmektedir. Ekstraksiyon sonrası elde edilen maternal ve hücre dışı fetal DNA (cffDNA), PCR aşamasında QF-PCR kiti kullanılarak spesifik STR markırları ile amplifiye edilir. Bu işlem, farklı kromozomlardaki belirli bölgelerin çoğaltılmasını sağlar.

Amplifikasyon ürünlerinin analizi, Applied Biosystems® ABI3500 kapiler elektroforez platformunda yüksek çözünürlüklü ayrışma ile gerçekleştirilmekte ve

GT500 (GeneScan™ 500 LIZ®) boy standardı kullanılarak kalibre edilmektedir. Multipleks floresan deteksiyon sistemi, eş zamanlı olarak birden fazla markırın analizine olanak tanımaktadır [39].

Amplifiye edilen ürünler, kapiler elektroforez temelli genetik analizör ile dizilenir ve fragment analizi gerçekleştirilir. Elde edilen veriler, biyoinformatik analiz sürecinde alel dozajı, trizomi ve diğer anöploidiler, allelik patern analizleri ile maternal ve paternal kalıtım açısından değerlendirilir. Son aşamada, analiz sonuçları trizomi 21, 13, 18, triploidi ve cinsiyet kromozom anöploidilerini içeren kapsamlı bir QF-PCR raporu halinde dokümante edilir. Bu akış, prenatal tanıda kromozomal anomalilerin moleküler düzeyde hızlı ve güvenilir tespitini mümkün kılmaktadır.




**Şekil 3.2** Prenatal tanıda QF-PCR test sürecinin numune alımından raporlamaya kadar olan analiz aşamaları, Biorender ile oluşturuldu.

Standartlaştırılmış primer-spesifik PCR protokolleri kullanılarak çoğaltılan DNA fragmentleri, Applied Biosystems® ABI3500 platformunda analiz edilir. Bu sistem, kapiler elektroforez teknolojisi ile yüksek çözünürlüklü ayrıştırma sağlarken, GT500 (GeneScan™ 500 LIZ®) boyut standardı ile hassas kalibrasyon gerçekleştirilir. Platform, multipleks floresan deteksiyon sistemi sayesinde birden fazla markırı aynı anda analiz edebilme kapasitesine sahiptir [39].

Veri analizi ve yorumlama süreci, GeneMapper Software 6 platformunda gerçekleştirilen analize dayanmaktadır. Bu süreçte alel spesifik floresan yoğunluk ölçümleri yapılmakta ve heterozigotluk/homozigotluk paternleri detaylı olarak değerlendirilmektedir. Metodolojinin geçerliliği, geniş bir popülasyon örnekleminde markır performans analizi ve popülasyonlar arası genetik değişkenlik değerlendirmesi ile sağlanmaktadır. Klinik uygulamada, bu metodoloji özellikle trizomi 13, 18, 21 ve sex kromozomları anöploidilerinin tespitinde yüksek

hassasiyet ve özgüllük göstermektedir. Genotip-fenotip korelasyon değerlendirmesi, klinik bulgularla genetik analiz sonuçlarının entegrasyonuna olanak tanımaktadır. Validasyon çalışmaları, metodun klinik tanı süreçlerinde güvenilir bir tanısal araç olarak kullanılabileceğini kanıtlamaktadır. Bu protokol, prenatal tanı süreçlerinde erken ve güvenilir genetik anomali tespitini mümkün kılarak, klinik karar verme mekanizmalarına önemli katkı sağlamaktadır. Multipleks markır analizi stratejisi, kromozomal anöploidilerin kapsamlı ve hızlı değerlendirilmesine olanak tanırken, metodun standardizasyonu ve validasyonu, klinik kullanım için sağlam bir temel oluşturur. Bu bağlamda, QF-PCR temelli anöploidi analizleri, modern prenatal tanı süreçlerinde güvenilir bir genetik inceleme yöntemi olarak yerini almıştır [39].

## 3.2 Kromozom Anomalileri

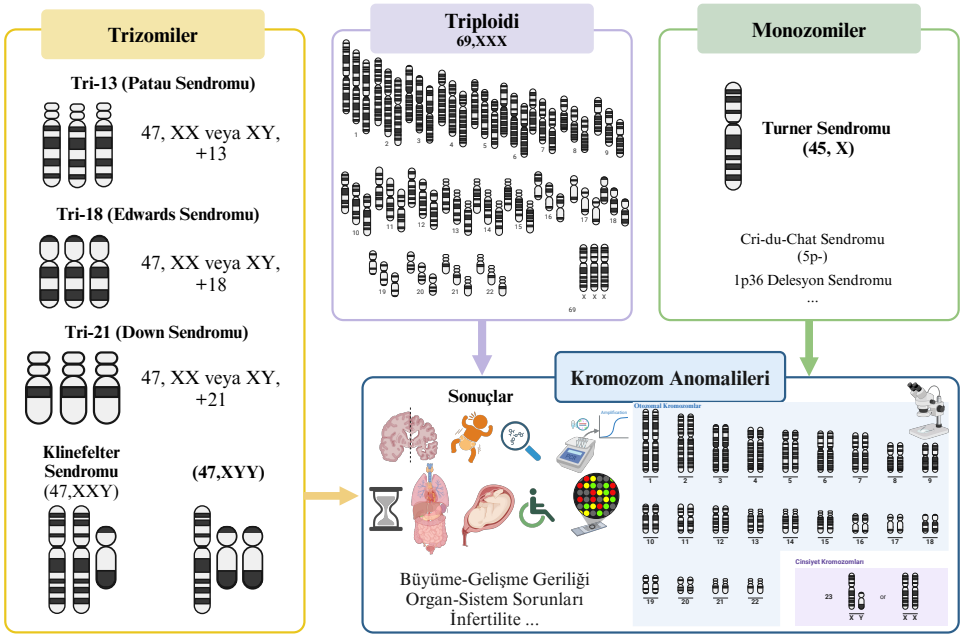
### 3.2.1 Anöploidiler

Anöploidi, bir organizmanın hücrelerinde kromozom sayısındaki değişiklikler ile tanımlanır. Normal insan kromozom sayısı 46 iken, anöploid durumlarda bu sayı artar veya azalır. Trizomi ve monozomi gibi anöploid durumları gösteren Şekil 3.3'de sol panelde trizomiler (Patau, Edwards, Down ve Klinefelter sendromları), orta panelde triploidiler ve sağ panelde monozomiler (Turner sendromu, Cri-du-Chat sendromu) gösterilmektedir. Alt panelde bu kromozomal anomalilerin fenotipik sonuçları ve karyotip analizi örnekleri yer almaktadır. Bunlar prenatal dönemde ciddi gelişimsel bozukluklar ve sağlık sorunlarına yol açabilir [11, 25].

Trizomiler, bir kromozomun üç kopyasının varlığı ile karakterize edilen kromozom anomalileridir. En sık görülen trizomiler arasında Trizomi 21 (Down sendromu), Trizomi 18 (Edwards sendromu) ve Trizomi 13 (Patau sendromu) bulunur. Bu durumlar, fetüste ciddi gelişimsel ve yapısal bozukluklara yol açabilir ve prenatal tanıda önemli bir yer tutar [9, 10].

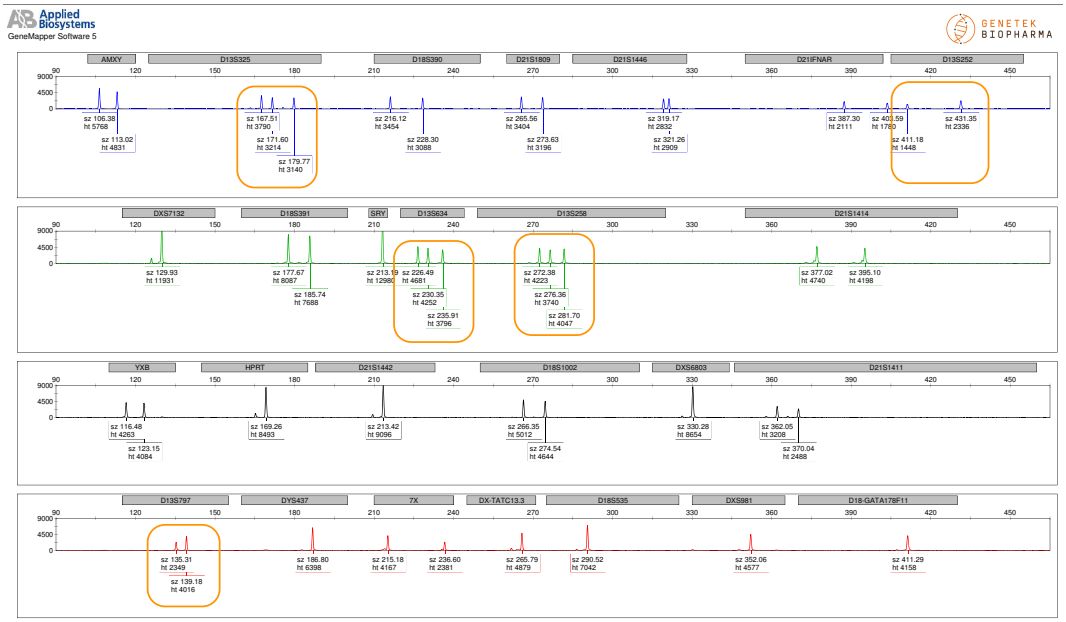
#### 3.2.1.1 Trizomi 13 (Patau Sendromu)

Trizomi 13, diğer adıyla Patau sendromu, 13. kromozomun fazladan bir kopyasının varlığına bağlı olarak ortaya çıkan, nadir görülen ancak ciddi bir kromozomal bozukluktur [12, 21]. Kalp kusurları, beyin anormallikleri ve yüz malformasyonları gibi birçok doğumsal anomalinin görüldüğü bu durum, genellikle bebeklerin çoğu birkaç haftadan fazla hayatta kalmaz [16]. Trizomi 13'ün prenatal tanısı, erken müdahale ve gebelik yönetimi açısından kritik öneme sahiptir ve QF-PCR, Array-CGH ve karyotipleme gibi çeşitli tekniklerle tespit edilebilir [4, 40]. Trizomi 13 vakalarının çoğu, gebeliğin ilk yarısında tespit edilip sonlandırılmaktadır; çünkü




**Şekil 3.3** Kromozomal anomalilerin sınıflandırılması ve klinik sonuçları: trizomi, triploidi ve monozomi sendromları, BioRender kullanılarak oluşturuldu.

bu fetüslerin yaşam beklentisi çok düşüktür [12, 21]. Trizomi 13, genellikle ultrason bulgularına dayanan rutin prenatal taramalar sırasında şiddetli klinik tablo nedeniyle erken tespit edilir [5].



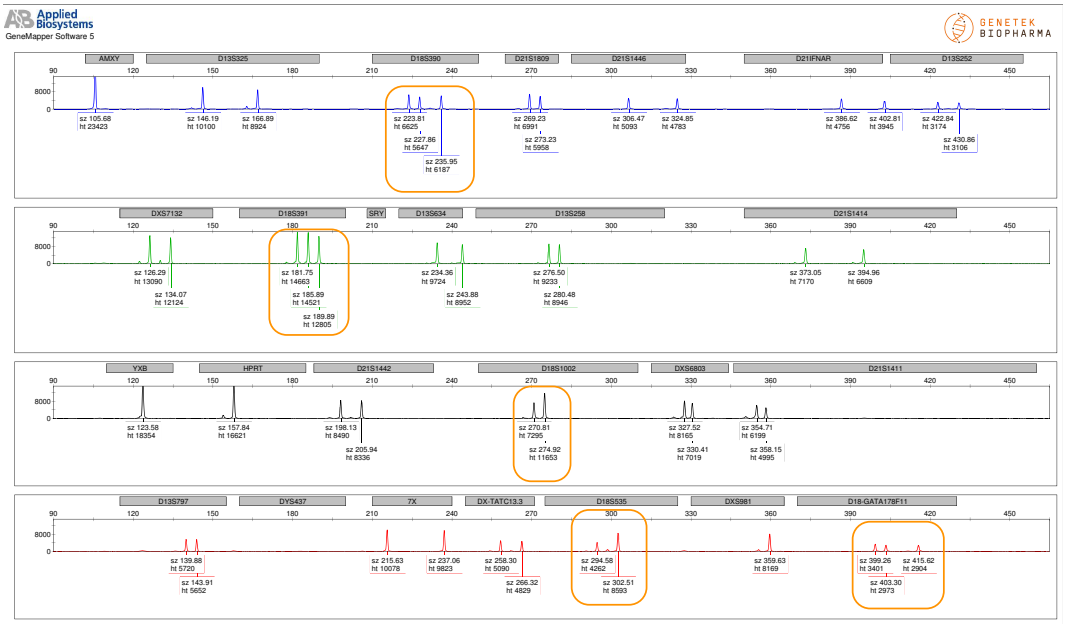

**Şekil 3.4** QF-PCR testi ile Trizomi 13 kromozom profil analizi [41].

GeneMapper Software 6 kullanılarak gerçekleştirilen QF-PCR elektroferogram analizi kromozom 13'e özgü STR markırlarını kullanarak kopya sayısı

değişikliklerini tespit eder. QF-PCR testi ile belirlenmiş Trizomi 13 profili Şekil 3.4'de gösterilmektedir. Turuncu kutular içinde işaretlenen bölgeler kromozom 13'e özgü STR markırlarında (D13S325, D13S252, D13S634, D13S258 ve D13S797) anormal kopya sayısını göstermektedir. Bu bölgelerde iki farklı patern gözlenmektedir: D13S325(167-171-179 bç), D13S634(226-230-235 bç) ve D13S258(272-276-281 bç) markırlarında triallelik patern (1:1:1 oranında üç eşit yükseklikte pik), D13S252(411-4312 bç) ve D13S797(135-1392 bç) markırlarında ise diallelik patern (1:2 oranında iki pik). Her iki patern de kromozom 13'ün üç kopya olduğunu göstererek Trizomi 13 tanısını desteklemektedir [41].

#### 3.2.1.2 Trizomi 18 (Edwards Sendromu)

Trizomi 18, fetüsün 18. kromozomunun üç kopyası ile oluşur ve ciddi fiziksel ve zihinsel bozukluklara neden olur. Edwards sendromuna sahip fetüslerde kalp defektleri, organ anomalileri ve kısa yaşam beklentisi gibi durumlar sıkça görülür. Prenatal tarama sırasında yüksek risk faktörleri saptandığında NIPT veya biyokimyasal testler önerilebilir; doğrulayıcı tanı için ise amniyosentez yapılır [23, 24, 27].



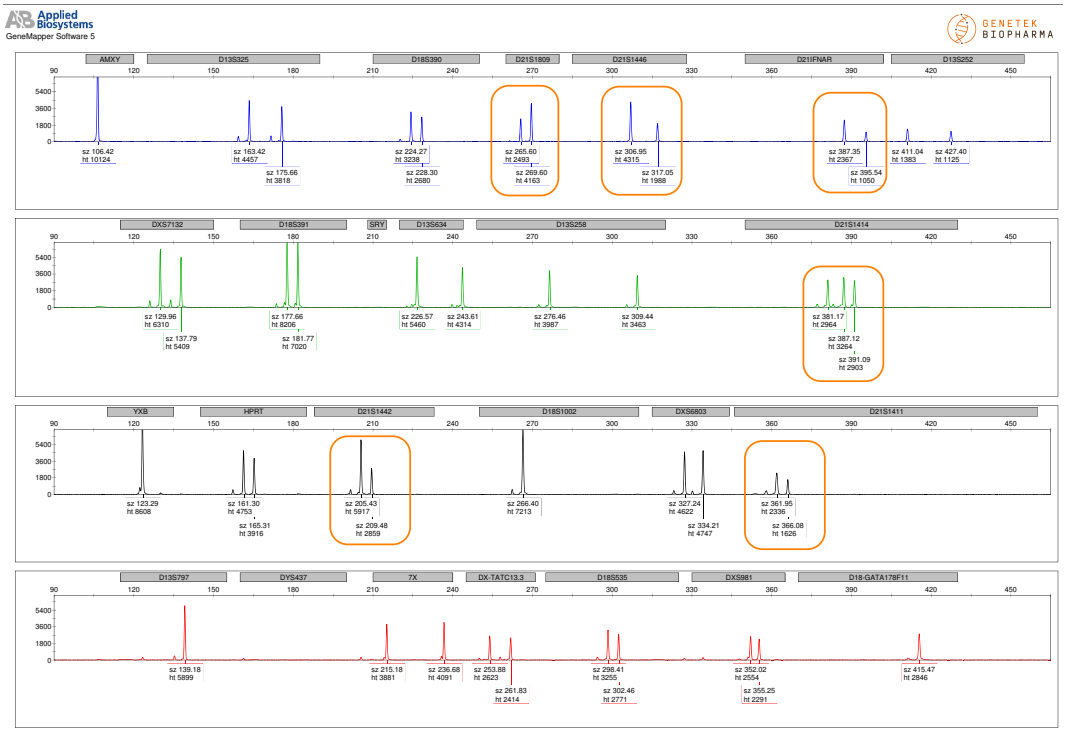

**Şekil 3.5** QF-PCR testi ile Trizomi 18 profil analizi [42]

QF-PCR elektroferogram analizi kromozom 18'e özgü STR markırlarını kullanarak kopya sayısı değişikliklerini tespit eder. QF-PCR testi ile belirlenmiş Trizomi 18 profili Şekil 3.5'de gösterilmektedir. Turuncu kutular içinde işaretlenen bölgeler kromozom 18'e özgü STR markırlarında (D18S390, D18S391, D18S1002, D18S535 ve D18-GATA178F11) anormal kopya sayısını göstermektedir. Bu bölgelerde iki farklı patern gözlenmektedir. D18S390(223-227-235 bç), D18S391(181-185-189

bç) ve D18-GATA178F11(399-403-415 bç) markırlarında triallelik patern (1:1:1 oranında üç eşit yükseklikte pik),D18S1002(270-274*2 bç) ve D18S535(294-302*2 bç) markırlarında ise diallelik patern (1:2 oranında iki pik). Her iki patern de kromozom 18'ün üç kopya olduğunu göstererek Trizomi 18 tanısını desteklemektedir [42].

#### 3.2.1.3 Trizomi 21 (Down Sendromu)

Trizomi 21, fetüsün 21. kromozomunun üç kopyasıyla karakterizedir ve en yaygın görülen trizomidir. QF-PCR elektroferogram analizi kromozom 21'e özgü STR markırlarını kullanarak kopya sayısı değişikliklerini tespit eder. QF-PCR testi ile belirlenmiş Trizomi 21 profili Şekil 3.6'de gösterilmektedir. Turuncu kutular içinde işaretlenen bölgeler kromozom 21'e özgü STR markırlarında D21S1809, D21S1446, D21IFNAR, D21S1414, D21S1442 ve D21S1411 anormal kopya sayısını göstermektedir. Bu bölgelerde iki farklı patern gözlenmektedir: D21S1414(381-387-391 bç) markırında triallelik patern (1:1:1 oranında üç eşit yükseklikte pik), D21S1809(265-2692 bç), D21S1446(3062-317 bç), D21IFNAR(3872-395 bç), D21S1442(2052-209 bç) ve D21S1411(361*2-366 bç) markırlarında ise diallelik patern (1:2 oranında iki pik). Her iki patern de kromozom 21'ün üç kopya olduğunu göstererek Trizomi 21 tanısını desteklemektedir [43].



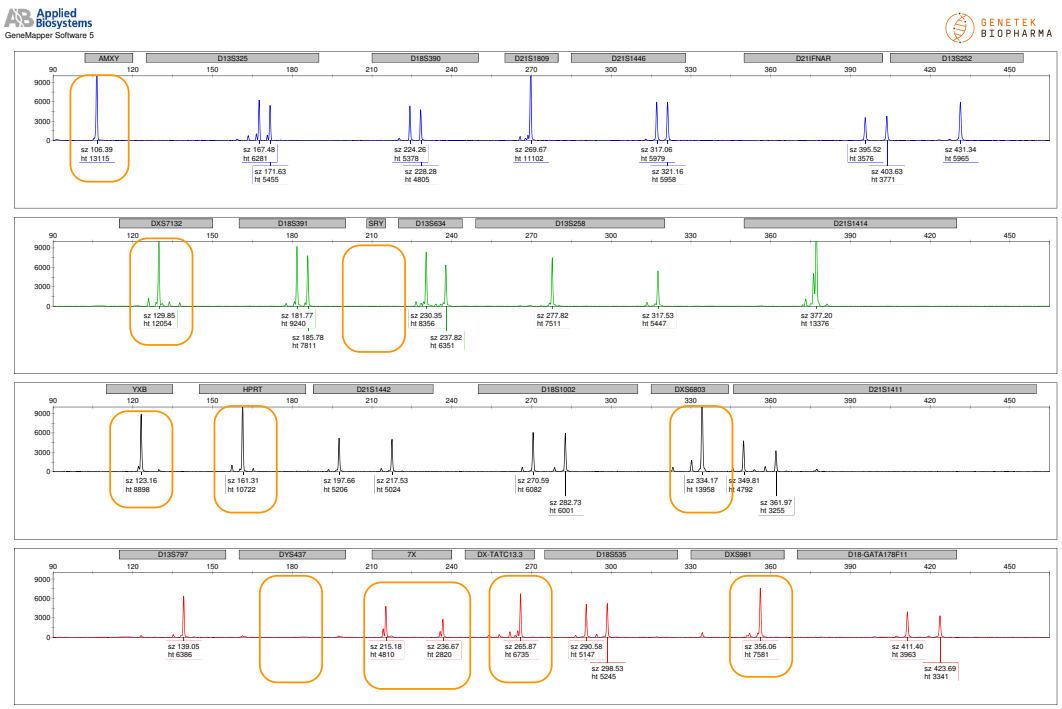

**Şekil 3.6** QF-PCR testi ile Trizomi 21 profil analizi [43]

Down sendromu, zihinsel ve fiziksel gelişim geriliğine, kalp anomalilerine, hipotoniye ve çeşitli sağlık sorunlarına yol açabilir. NIPT, Down sendromunun

taranmasında yüksek hassasiyet sunar ve doğrulayıcı tanı için amniyosentez veya CVS gibi invaziv testlerle desteklenir [10, 11].

#### 3.2.1.4 Monozomi

Monozomi, bir kromozomun tek kopyasının bulunması durumudur ve genellikle ölümcül sonuçlar doğurur. En bilinen örneği, cinsiyet kromozomlarında görülen Turner sendromudur. Turner sendromunda fetüs, X kromozomunun bir kopyasına sahipken ikinci X veya Y kromozomu eksiktir. Bu durum, kısa boy, infertilite ve kalp problemleri gibi çeşitli sağlık sorunlarına yol açabilir [12, 14]. QF-PCR testi ile belirlenmiş Turner sendromu profili Şekil 3.7'de turuncu kutularda işaretlenen bölgelerde tek pik gözlenmektedir. İşaretli markırlarda monoallelik patern (tek pik) görülmesi, X kromozomunun tek kopya olduğunu göstererek Turner Sendromu tanısını desteklemektedir. Bu patern 45,X karyotipini işaret etmektedir [44].

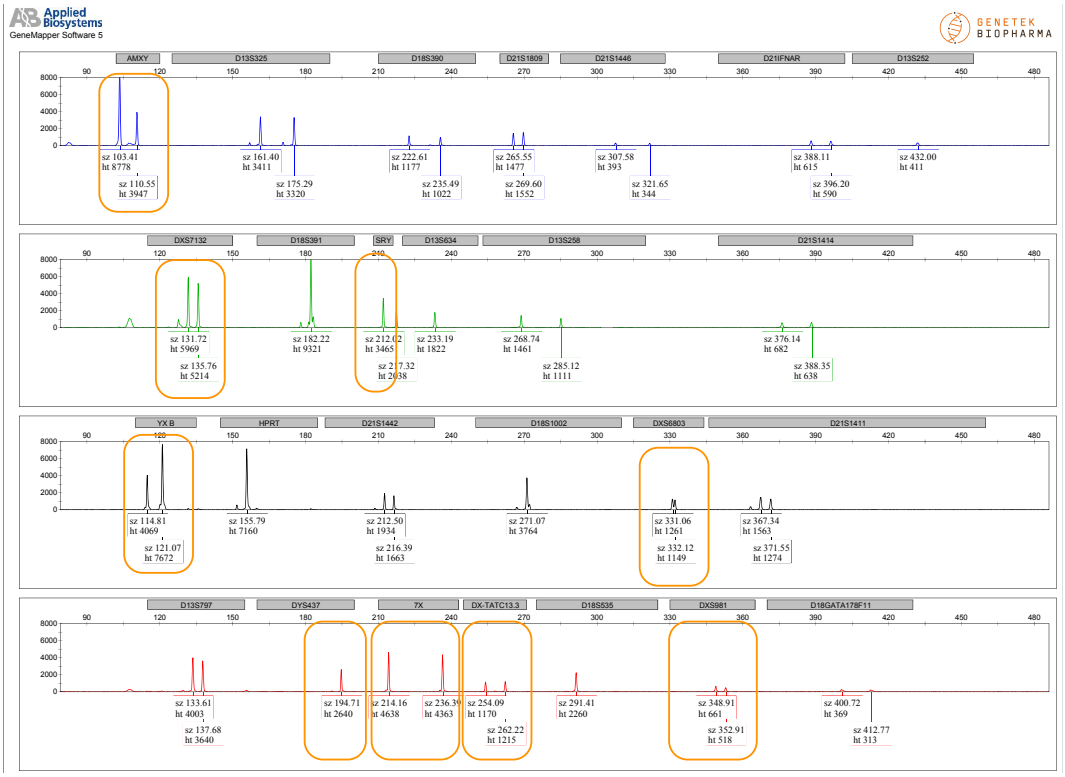



**Şekil 3.7** QF-PCR testi ile Turner Sendromu profil analizi [44].

Mozaiklik, fetüste birden fazla hücre tipi veya kromozomal yapı bulunması anlamına gelir. Mozaik durumlar, tüm hücrelerde aynı anomaliyi göstermediği için tanı ve değerlendirmede zorluk yaratabilir. Mozaiklik durumlarında genetik danışmanlık, sonuçların yorumlanmasında önemli bir rol oynar çünkü bu durum, fenotipik varyasyonlara neden olabilir ve klinik tablonun ciddiyetini tahmin etmeyi zorlaştırır [10, 15].

### 3.2.1.5 Klinefelter Sendromu (XXY) ve Triple X Sendromu (XXX)

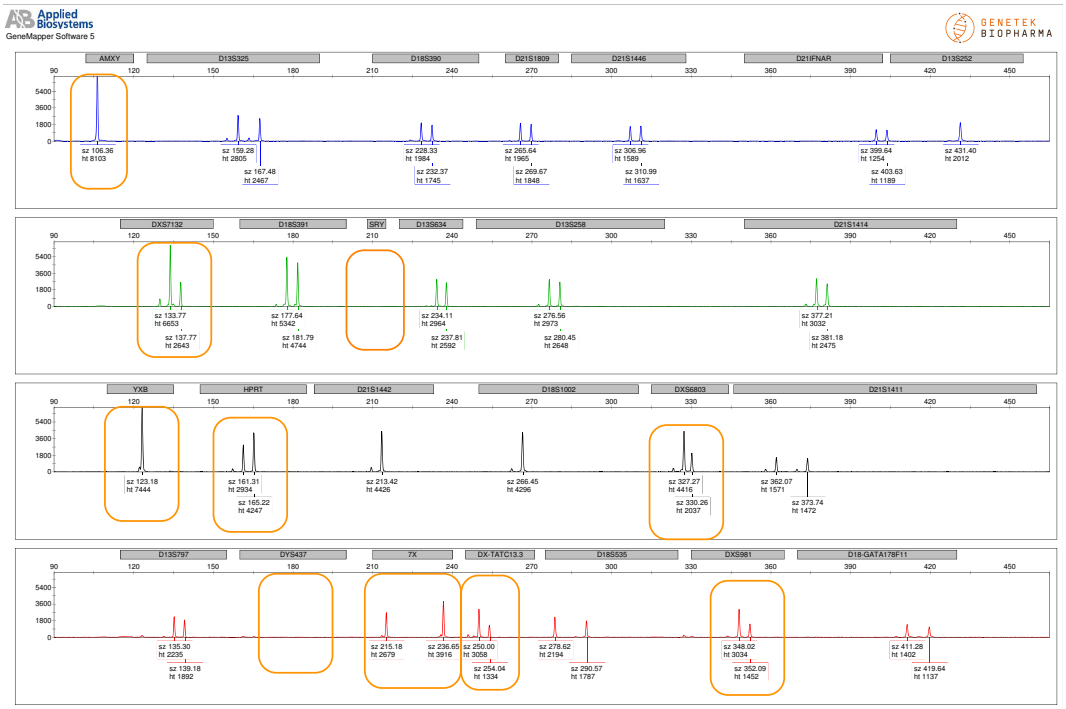
Klinefelter sendromu (XXY) ve Triple X sendromu (XXX), cinsiyet kromozomu anöploidileri (SCA) olarak adlandırılan genetik anomaliler grubuna aittir. Her iki sendrom da kromozom sayısındaki sapmalar nedeniyle ortaya çıkar ve çeşitli klinik, nörolojik ve psikolojik etkilerle ilişkilidir [45, 46].




**Şekil 3.8** QF-PCR testi ile Klinefelter Sendromu profil analizi [47].

Klinefelter Sendromu (XXY), genetik olarak erkek bireylerde ek bir X kromozomunun bulunmasıyla karakterizedir. GeneMapper Software 6 kullanılarak gerçekleştirilen QF-PCR elektroferogram analizi Şekil 3.8'de turuncu kutularda işaretlenen bölgelerde cinsiyet kromozom markırlarında AMXY(103*2-110 bç) ve YXB(114-121*2 bç) diallelik patern (2:1 oranında iki pik) gözlenmektedir. X kromozomuna özgü markırlarda iki kopya, Y kromozomuna özgü markırlarda tek kopya olduğunu gösteren bu paternler, 47,XXY karyotipini işaret ederek Klinefelter Sendromu tanısını desteklemektedir [47]. Cinsiyet kromozomlarına özgü AMXY ve YXB STR markırlarındaki 2:1 oranında iki pik durumu, çoğunlukla erkeklerde infertilite, testiküler az gelişim ve azalmış testosteron seviyeleri gibi endokrin bozukluklara neden olur. Ayrıca, Klinefelter sendromu olan bireylerde öğrenme güçlükleri, dikkat eksikliği hiperaktivite bozukluğu (DEHB) ve otizm spektrum bozukluğu gibi nörolojik sorunların görülme sıklığı artmıştır. Erken tanı ve hormon tedavileri, özellikle ergenlik dönemindeki testosteron tedavileri, fiziksel ve psikolojik gelişimi destekleyebilir ve bireylerin yaşam kalitesini artırabilir [46, 48].

Triple X Sendromu (XXX) QF-PCR analizi Şekil 3.9'de turuncu kutularda gösterilmektedir. QF-PCR elektroferogram analizindeki AMXY ve YXB, STR makerlarında işaretlenen bölgelerde X kromozomuna özgü markırlarda monoallelik patern (tek pik) gözlenmektedir. Bu patern, X kromozomunun varlığını gösterirken, SRY markırında Y kromozomunda sinyal görülmemesi 47,XXX karyotipini işaret etmekte ve Triple X Sendromu tanısını desteklemektedir [49]. Kadınlarda fazladan bir X kromozomunun bulunmasıyla ortaya çıkar ve genellikle yaşam boyunca hafif fiziksel semptomlarla seyredebilir. Bununla birlikte, öğrenme güçlükleri, hafif motor beceri bozuklukları ve sosyal iletişimde zorluklar gibi nöropsikiyatrik belirtiler gözlemlenebilir. Triple X sendromlu bireylerde genellikle doğurganlık korunur, ancak bazı vakalarda hormonal dengesizlikler ve üreme sistemi ile ilgili komplikasyonlar görülebilir. Bu sendrom, genellikle mozaiklik gibi değişken karyotiplerle de kendini gösterebilir ve çoğu zaman teşhis edilmeden kalır [46, 50].

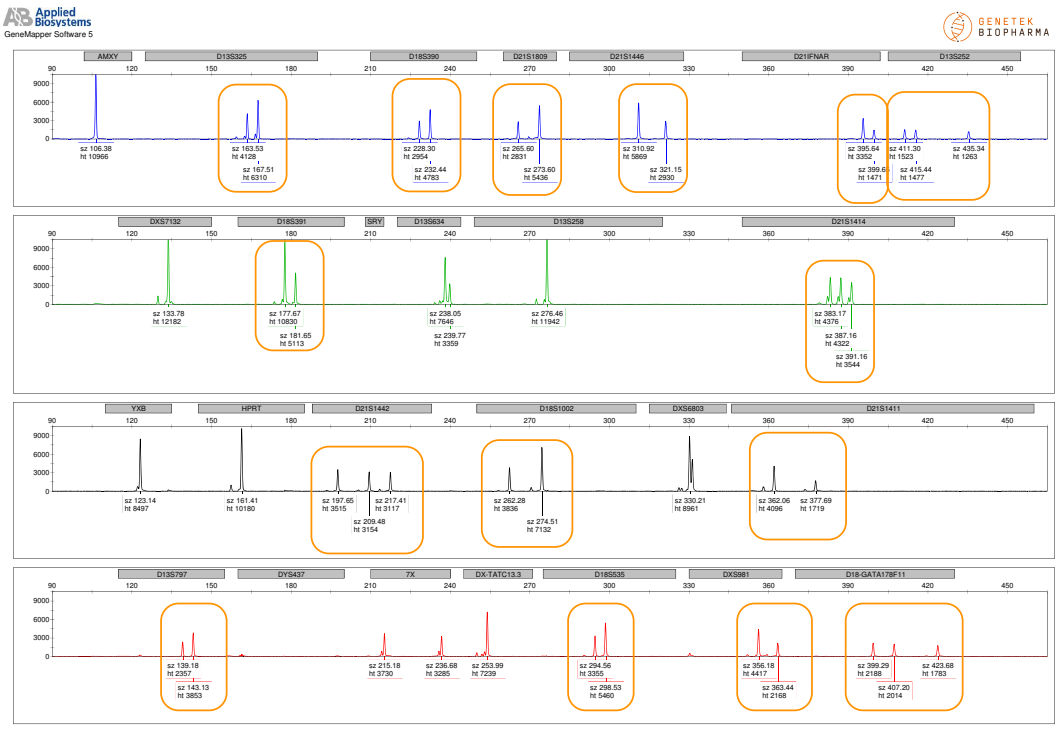



**Şekil 3.9** QF-PCR testi ile Triple X profil analizi [49].

Her iki sendrom da genetik danışmanlık ve uygun tedavi yaklaşımlarını gerektirir. Klinefelter sendromunda testosteron tedavisi, fiziksel ve zihinsel gelişimi desteklemek için kritik bir rol oynar. Triple X sendromunda ise erken eğitimsel destek ve psikososyal müdahaleler, bireylerin sosyal ve akademik başarılarını artırabilir. Bu iki durum, cinsiyet kromozomlarının nörolojik ve psikiyatrik fenotipler üzerindeki etkilerini anlamak için önemli modeller sunar ve bireysel hasta yönetimi stratejilerini yönlendirmek için genetik faktörlerin incelenmesinde temel oluşturmaktadır [48, 50].

#### 3.2.1.6 Triploidi

Triploidi, bir organizmanın her hücresinde üç set kromozom bulunması durumu olup, normal diploid yapıya sahip bir insanın 46 kromozomuna karşılık 69 kromozom taşıdığı anlamına gelir. Bu genetik anomali, embriyonun her bir kromozomdan üç kopya taşıması sonucunda ortaya çıkar ve ciddi sağlık sorunlarına neden olur. Triploidi, prenatal tanıda karşılaşılan nadir ancak ciddi bir durumdur ve çoğunlukla yaşamla bağdaşmaz. Triploidi vakalarının büyük bir çoğunluğu düşükle sonuçlanır ya da doğumdan kısa bir süre sonra bebek hayatını kaybeder [10, 23, 24]. Bu anomali, embriyonun genetik materyalinin kökenine bağlı olarak "diandrik" (baba kökenli) veya "digynik" (anne kökenli) olarak sınıflandırılır ve bu terimler, literatürde doğru ve yaygın kabul gören kavramlardır [51, 52].




**Şekil 3.10** QF-PCR testi ile Triploidi profil analizi [53].

Diandrik triploidi, ekstra haploid kromozom setinin babadan kaynaklandığı vakaları tanımlar. Genellikle dispermi, yani bir oositin iki sperm tarafından döllenmesi sonucu oluşur ve bu mekanizma diandrik vakaların büyük çoğunluğundan sorumludur. Daha nadir durumlarda, diploid bir spermin döllenmeye katılması ile ortaya çıkar. Diandrik triploidinin klinik görünümü, görece iyi gelişmiş fetüsler, kistik plasental değişiklikleri ve trofoblast hiperplazisi ile karakterizedir. Bununla birlikte, preeklampsi ve gestasyonel trofoblastik neoplazi gibi maternal komplikasyon riskleri oldukça yüksektir [52, 54]. Buna karşılık, digynik triploidi ekstra kromozom setinin anneden geldiği vakaları ifade eder. Bu durum, genellikle bir oositin mayoz I veya II sırasında oluşan hatalar sonucu diploid hale gelmesi ya da endoreduplikasyon

(polisiklik genom amplifikasyonu süreci) olarak bilinen DNA replike hatası ile oluşur. Digynik triploidi, fetüste asimetrik büyüme geriliği, relatif makrosefali ve non-kistik plasenta gibi özelliklerle kendini gösterir. Anneye yönelik komplikasyon riski, diandrik triploidiye göre daha düşük olsa da, gebelik sonuçları genellikle olumsuzdur [54, 55].

Prenatal tanıda triploidi teşhisi ve parental kökenin belirlenmesi, hem maternal komplikasyonları öngörmek hem de gebelik yönetimini yönlendirmek açısından kritik öneme sahiptir. Ultrasonografi, plasental ve fetal anomalilerin değerlendirilmesi için ilk tanı yöntemlerinden biri olsa da, özellikle parental kökenin saptanmasında moleküler yöntemler daha yüksek hassasiyet sunar. QF-PCR testi ile belirlenmiş triploidi profili Şekil 3.10'de turuncu kutularda işaretlenen bölgelerde incelenen tüm kromozomal markırlarda triallelik patern (1:1:1 oranında üç eşit yükseklikte pik) veya diallelik patern (2:1 oranında farklı yükseklikte iki pik) gözlenmektedir. Bu paternler, tüm kromozomların üç kopya olduğunu göstererek 69,XXY veya 69,XXX karyotipini işaret etmekte ve Triploidi tanısını desteklemektedir [53]. Bu yöntem, klasik histopatolojik incelemelere kıyasla daha doğru sonuçlar sağlar ve diandrik triploidi ile ilişkili parsiyel hidatidiform mol (kısmi mol gebelik) tanısında klinik uygulamaya değerli bir katkı sunar [9]. Histopatolojik değerlendirme ise özellikle plasental yapıların karakterize edilmesinde yardımcı bir yöntem olarak kullanılabilir [52, 54].

## 3.3 Kromozomal Anöploidi Tespitinde Diğer Tanısal Yöntemler

Günümüzde genetik hastalıkların prenatal tanısında ilerleyen teknoloji ile gelişmeler kaydedilmektedir. Down sendromu ve diğer kromozom bozuklukları gibi genetik hastalıkların teşhisinde genler, miRNA'lar ve moleküler yöntemler önemli bir rol oynamaktadır. Floresan in situ hibridizasyon (FISH) ve kantitatif floresan polimeraz zincir reaksiyonu (QF-PCR) gibi hızlı moleküler yöntemler, prenatal teşhis süreçlerinde yaygın olarak kullanılmakta olup, doğru ve hızlı sonuçlar sunmaktadır. Bununla beraber Array-CGH, tam genom dizilimi (WGS) ve kromozomal mikrodizin analizi (CMA) gibi yöntemler, konjenital anomalileri olan fetüsler ve bebeklerde genetik varyantların tespitinde kullanılmakta ve bu yöntemlerin ek katkıları, tanısal süreçlerdeki önemini artırmaktadır. Ancak, her bir yöntemin kendine özgü avantajları ve kısıtlamaları mevcuttur, bu da farklı vakalar için yöntem seçiminde dikkatli bir değerlendirme gerektirmektedir. Çalışmalarda QF-PCR'ın önemi yadsınamaz. QF-PCR/CMA(kromozomal mikrodizin analizine)'nin ek tanısal getirisi tüm vakalar için WGS(tam genom diziliminin)'ye göre fazladır [56].

### 3.3.1 Array-CGH

Karşılaştırmalı Genomik Hibridizasyon (Array-CGH-Array Comparative Genomic Hybridization), prenatal tanıda kromozomal dengesizlikleri, silinmeleri ve duplikasyonları tüm genom boyunca tarayan ve tespit etmek için kullanılan yüksek çözünürlüklü bir moleküler tekniktir [16, 57]. Geleneksel karyotiplemeye kıyasla daha yüksek bir çözünürlük sunar ve fark edilmeyecek olan mikroskobik kromozomal anormalliklerin tespit edilmesine olanak tanır [16]. Bu teknik, hasta DNA'sı ile referans DNA'nın karşılaştırılmasına dayanır ve genomik dengesizlikleri tespit edebilir. Oligonükleotid problar kullanarak kopya sayısı varyasyonları (CNV) tespiti gerçekleştirebilmektedir [57]. Array-CGH hücre kültürü gerektirmez, bu da sonuç almak için gereken süreyi kısaltır ve yüksek riskli gebeliklerde ebeveynlerin kaygısını azaltır [40, 58]. Array-CGH'nin kullanımı, özellikle gelişimsel gecikmeler ve doğuştan gelen malformasyonlarla (doğumsal anomaliler) ilişkili durumları tanımlamak için daha ayrıntılı genetik bilgilere ihtiyaç duyulan vakalarda giderek daha önemli hale gelmiştir [59].

### 3.3.2 Kromozomal Mikroarray Analizi (CMA)

CMA (Kromozomal Mikroarray Analizi), prenatal tanıda ve doğum sonrası genetik hastalıkların tanısında kullanılan yüksek çözünürlüklü bir moleküler sitogenetik testtir. CMA, genom boyunca CNV ve SNP varyant analizlerini gerçekleştirebilmektedir. Özellikle prenatal tanıda ultrason bulgularında anomali saptanan fetüslerde, konjenital anomalilerin genetik nedenlerini araştırmada, düşük materyallerinin incelenmesinde ve prenatal tarama testlerinde pozitif sonuçların doğrulanmasında kullanımı önerilmektedir [57].

CMA'nın tanısal verimi, standart karyotip analizine göre daha yüksektir ve yaklaşık %5-10 ek tanı sağlamaktadır [60]. Prenatal tanıda CMA'nın tanısal verimi %7.8, postnatal vakalarda ise %16.3 olarak rapor edilmiştir [57]. Ayrıca kültür gerektirmediği için sonuç süresi daha kısadır ve mozaisizm durumlarında daha hassas sonuç vermektedir. CMA, 5-400 kb arasında çözünürlüğe sahip olup, platformlar arasında farklılıklar gösterebilmektedir. CMA teknolojisi, >%10-15 mozaik oranını tespit edebilmekte ve uniparental dizomi (UPD) analizine de olanak sağlamaktadır. SNP array teknolojisi ile birlikte kullanıldığında, sadece hasta örneği yeterli olabilmektedir. Bu özellikleri ile CMA, prenatal tanı alanında kompleks genetik araştırmalar, sendromik olgular ve Mendelian hastalık şüphesi olan durumlarda tercih edilmektedir. Bununla birlikte, maliyet etkinliği açısından değerlendirildiğinde daha yüksek maliyetli bir test olup, veri analiz kompleksitesi de yüksek düzeydedir [56, 57].

Prenatal tanıda kullanılan ve daha kapsamlı genomik veri elde edilmesini sağlayan Array-CGH ve CMA teknolojilerinın karşılaştırılması Tablo 3.2'de verilmiştir. Array-CGH kopya sayısı varyasyonlarını (CNV) tespit ederken, CMA buna ek olarak SNP varyant analizlerini de gerçekleştirebilmektedir. Prenatal tanıda, QF-PCR'ın duyarlılığı %99.21, özgüllüğü %99.95 olarak bildirilmiştir [57]. Array-CGH için tek örnek veya referans DNA gerekirken, CMA'da sadece hasta örneği yeterlidir. CMA'nın tanısal verimi, özellikle ultrason anomalisi olan fetüslerde %7.8, postnatal vakalarda ise %16.3 olarak rapor edilmiştir [57]. Her iki yöntemin de kullanım alanları farklıdır. Array-CGH daha çok postnatal CNV analizi, kanser genomik profili ve hızlı preliminer sonuç gereken durumlarda tercih edilirken; CMA prenatal tanı, sendromik olgular, Mendelian hastalık şüphesi ve kompleks genetik araştırmalarda tercih edilmektedir.

**Tablo 3.2** Array-CGH ve CMA karşılaştırması [56, 57, 60]

| **Özellik** | **Array-CGH** | **CMA** |
|---|---|---|
| Temel Analiz Kapasitesi | Kopya sayısı varyasyonları (CNV) tespiti | CNV + SNP varyant analizi |
| Platform Teknolojisi | Oligonükleotid problar ile CNV tespiti | SNP probları veya kombine (CNV+SNP) problar |
| Genomik Çözünürlük | 10-400 kb (platform bağımlı) | 5-400 kb (platform bağımlı) |
| Mozaizm Saptama Limiti | >%20-30 mozaik oranı | >%10-15 mozaik oranı |
| UPD Analizi | Tespit edilemez | Tespit edilebilir |
| ROH/AOH Analizi | Tespit edilemez | Tespit edilebilir |
| Örnek Gereklilikleri | Tek örnek veya referans DNA | Sadece hasta örneği yeterli |
| Optimum Kullanım Alanları | • Postnatal CNV analizi<br>• Kanser genomik profili<br>• Hızlı preliminer sonuç gereken durumlar | • Prenatal tanı<br>• Sendromik olgular<br>• Mendelian hastalık şüphesi<br>• Kompleks genetik araştırmalar |
| Maliyet Etkinliği | Daha ekonomik | Daha yüksek maliyet |
| Veri Analiz Kompleksitesi | Orta düzey | Yüksek düzey |

### 3.3.3 Karyotipleme

Karyotipleme, özellikle anöploidi, translokasyonlar ve büyük yapısal yeniden düzenlemeler gibi kromozomal anormallikleri tespit etme konusunda uzun zamandır prenatal tanının altın standardı olarak kabul edilmektedir [3, 4]. Bu teknik, amniyosentez veya koryon villus biyopsisi yoluyla alınan fetal hücrelerin kültüre edilmesini ve ardından metafaz kromozomlarının mikroskop altında boyanarak analiz edilmesini içerir [16]. Oldukça güvenilir olmasına rağmen, karyotipleme zaman

alıcıdır; sonuçlar genellikle 10 ya da 14 gün sürer ve çözünürlüğü sınırlıdır, bu da 5 megabazdan küçük anormalliklerin tespit edilmesini zorlaştırır [6]. QF-PCR ve Array-CGH gibi daha hızlı tekniklerin benimsenmesiyle birlikte, karyotiplemenin rolü giderek evrilmekte olsa da dengeli kromozomal yeniden düzenlemeleri tespit etmek için vazgeçilmez bir yöntem olarak kalmaktadır [3, 59].

### 3.3.4 FISH-Floresan In Situ Hibridizasyon

Floresan in situ hibridizasyon (FISH), kromozomal anomalilerin hızlı ve doğru tespitinde kullanılan önemli bir moleküler sitogenetik tekniktir. FISH analizi, floresan işaretli DNA problarının hedef kromozom bölgelerine spesifik olarak hibridize olması prensibine dayanır [61]. Özellikle prenatal tanıda, kromozomal anöploidilerin ve mozaisizmin belirlenmesinde yaygın olarak kullanılmaktadır. FISH metodolojisinde, genetik materyalin kromozomal düzeyde analizi için metafaz veya interfaz hücrelerinden yararlanılır. Örneğin, Down sendromu şüphesi olan vakalarda 21. kromozom için spesifik problar kullanılarak trizomi 21 varlığı araştırılabilir [62]. FISH tekniği aynı zamanda yapısal kromozom anomalilerinin, kırık noktalarının ve translokasyonların belirlenmesinde de kullanılabilmektedir.

# 4
# MAKİNE ÖĞRENMESİ ALGORİTMALARI

## 4.1 Makine Öğrenmesi Algoritmaları

### 4.1.1 Lojistik Regresyon (Logistic Regression) Algoritması

Lojistik Regresyon Algoritması, sınıflandırma problemlerinde sıkça tercih edilen bir istatistiksel makine öğrenimi algoritmasıdır. Özellikle ikili sınıflandırmalarda güçlü bir performans sergiler ve doğrusal ilişkilerde yüksek doğruluk sağlar. Lojistik Regresyon, özellikle kanser semptomları gibi hastalık verilerinde belirleyici faktörlerin anlaşılmasına olanak tanır. Zeinali ve arkadaşları, bu algoritmayı kanser semptomlarını öngörmek için kullanmış ve yaş, cinsiyet gibi temel değişkenlerin belirleyiciliğini göstermiştir. Bu tür sınıflandırmalar, hasta bireyselleştirilmiş bakım stratejilerinin geliştirilmesine önemli katkılar sunmaktadır [63].

Lojistik Regresyon Algoritması'nın sağlık alanındaki uygulamalarında öne çıkan avantajları, klinik karar verme süreçlerine sağladığı önemli katkılarla dikkat çekmektedir. Model çıktılarının katsayılar aracılığıyla doğrudan yorumlanabilir olması, klinik personelin analiz sonuçlarını etkili bir şekilde değerlendirebilmesine olanak tanımaktadır [64]. Özellikle hesaplama maliyetinin düşük olması ve küçük veri setlerinde gösterdiği yüksek performans, klinik uygulamalarda hızlı geri bildirim gerektiren durumlar için ideal bir çözüm sunmaktadır [65]. Buna ek olarak, doğrusal sınıflandırma problemlerinde gösterdiği yüksek başarı oranı, özellikle hasta verilerinde belirli bir doğrusal ayrım olduğu durumlarda etkin sonuçlar elde edilmesini sağlamaktadır [66].

Bununla birlikte, Lojistik Regresyon Algoritması'nın bazı sınırlamaları ve dezavantajları da bulunmaktadır. Özellikle doğrusal olmayan karmaşık ilişkileri içeren veri setlerinde modelin performansı önemli ölçüde düşmektedir. Nörodejeneratif hastalıklar üzerine yapılan araştırmalarda, Przybyszewski ve arkadaşları tarafından belirtildiği gibi, model karmaşık veri yapılarını analiz etmekte yetersiz kalabilmektedir [66]. Ayrıca, küçük ölçekli veya aşırı gürültülü veri setlerinde yetersiz kalmasından dolayı modelin performansı önemli ölçüde düşmekte,

bu da klinik uygulamalarda güvenilirlik sorunlarına yol açabilmektedir [67].

### 4.1.2 Karar Ağacı (Decision Tree) Algoritması

Karar Ağacı Algoritması, verileri dallara ayırarak karar verme sürecini daha anlaşılır hale getiren bir algoritmadır. Her bir düğümde belirli bir özniteliğe göre veri dallanır ve en sonunda sınıflandırma veya regresyon sonucuna ulaşılır. Sivajohan ve arkadaşlarının endometriozis çalışmasında, Karar Ağacı Algoritması ile bu hastalığın tanısında doğruluğun artırılması hedeflenmiştir. Bu algoritma, klinik verilere dayalı tanı süreçlerinde tercih edilmektedir [67].

Karar Ağacı Algoritması, veri analizi ve sınıflandırma problemlerinde önemli avantajlar sunan bir yöntemdir. Bu algoritmanın en dikkat çekici özelliklerinden biri, kullanıcıların veriyi daha kolay anlamasını sağlaması ve algoritmanın çıktılarını görselleştirerek değişkenlerin etkisini açıkça göstermesidir. Özellikle Random Forest gibi ensemble yöntemlerle karşılaştırıldığında, Karar Ağacı'nın daha kolay anlaşılabilir ve yorumlanabilir sonuçlar sunduğu ortaya konmuştur [64]. Öznitelik seçiminde, modelin hangi özniteliklerin daha etkili olduğunu göstermesi önemli bir avantaj sağlamakta ve verinin farklı özelliklerini belirleyerek modelin genel performansını artırabilmektedir [67]. Ayrıca, özellikle küçük veri setlerinde hızlı sonuç vermesi ve düşük hesaplama süresi, klinik verilerde pratik bir çözüm olarak kullanılmasını sağlamaktadır [68].

Bununla birlikte, Karar Ağacı Algoritması'nın bazı önemli sınırlamaları ve dezavantajları bulunmaktadır. En önemli dezavantajlarından biri, modelin aşırı öğrenme (overfitting) eğiliminde olmasıdır. Bu durum, veri setindeki her ayrıntıyı öğrenmeye çalışmasına neden olabilmekte ve dengesiz veri setlerinde yetersiz kalmasına yol açabilmektedir [69]. Ayrıca, küçük veri değişimlerinden bile etkilenebilmesi ve model performansının dalgalanabilmesi, her yeni veri setinde farklı performansların gözlenmesine neden olmaktadır. Bu hassasiyet, özellikle klinik sonuçların doğruluğunu olumsuz yönde etkileyebilmektedir [65].

### 4.1.3 Random Forest (Rastgele Orman) Algoritması

Random Forest Algoritması, çok sayıda karar ağacını birleştirerek daha yüksek doğruluk sağlayan bir ensemble öğrenme yöntemi ve sınıflandırma problemlerinde yüksek doğruluk oranları sunar. Her ağaç, veri setinden rastgele örneklerle oluşturulur ve modellerin birleşimi ile daha sağlam bir sonuç elde edilir. Vakitbilir ve arkadaşlarının travmatik beyin hasarı çalışmasında Random Forest, intrakraniyal basınç (ICP, intracranial pressure) gibi beyin fizyolojik parametrelerinin tahmininde

üstün performans sergileyerek, Random Forest'ın sağlık alanındaki potansiyelini ortaya koymaktadır [69].

Random Forest Algoritması, makine öğrenmesi alanında güçlü ve güvenilir bir yöntem olarak öne çıkmaktadır. Bu algoritmanın en önemli avantajlarından biri, aşırı öğrenme sorununu büyük ölçüde azaltması ve farklı karar ağaçlarının birleşimi ile genelleme kapasitesini artırmasıdır. Sağlık verilerinde değişken önemini inceleme çalışmalarında, veriyi aşırı öğrenmeden modelleme kapasitesinin yüksek olduğu görülmüştür [64]. Özellikle dengesiz veri setlerinde performans açısından avantajlı olan bu algoritma, kanser semptomlarını tahmin etme çalışmalarında doğruluk ve hassasiyet açısından öne çıkmıştır [63]. Ayrıca, hangi özelliklerin sınıflandırmada veya regresyonda daha etkili olduğunu belirleyerek, kullanıcıya özniteliklerin önemi konusunda değerli bilgiler sunmaktadır [67].

Bununla birlikte, Random Forest algoritmasının bazı önemli dezavantajları da bulunmaktadır. Birçok ağacın birleşmesinden kaynaklanan karmaşık yapısı nedeniyle, hesaplama maliyeti oldukça yüksektir ve bu durum özellikle büyük veri setlerinde algoritmanın yavaş çalışmasına neden olabilmektedir [69]. Ayrıca, karar ağaçlarının birleşiminden oluşan yapısı nedeniyle modelin iç işleyişini anlamak oldukça zordur. Bu "kara kutu" özelliği, kararların şeffaflığını azaltmakta ve modelin doğruluğunu yorumlama konusunda önemli sınırlamalar yaratmaktadır [68].

### 4.1.4 K-Nearest Neighbor (KNN, K-En Yakın Komşu) Algoritması

KNN algoritması, verileri en yakın komşularının özelliklerine göre sınıflandıran veya regresyon yapan denetimli bir öğrenme algoritmasıdır. Yeni bir veri noktası, kendisine en yakın K adet komşusunun sınıf etiketlerine bakılarak sınıflandırılır. K parametresi, modelin önemli bir hiperparametresidir ve optimal değeri çapraz doğrulama yöntemiyle belirlenir. En iyi performansı elde etmek için farklı K değerleri test edilir ve veri setine en uygun K değeri seçilir. Bu algoritma, özellikle örüntü tanıma ve sınıflandırma problemlerinde sıklıkla kullanılır. Amin ve arkadaşlarının antibiyotik reçeteleme çalışmasında KNN, gereksiz reçetelerin azaltılmasına katkı sağlamıştır ve bu tür halk sağlığı sorunlarının çözümünde umut vadetmektedir [70].

Algoritmanın mantığının basit olması ve sınıflandırma kararlarının kolayca anlaşılabilir olması, özellikle ilk analizlerde ve küçük veri setlerinde tercih edilmesini sağlamaktadır [63]. Hem sınıflandırma hem de regresyon problemlerinde kullanılabilen bu algoritma, sağlık verilerinde farklı sınıflar arasındaki ilişkileri belirlemede güçlü bir model olarak öne çıkmaktadır [65]. Veri dağılımına dair

herhangi bir varsayım yapmaması nedeniyle parametrik olmayan veri setlerinde başarılı sonuçlar vermektedir [67].

Öte yandan, KNN algoritmasının bazı önemli dezavantajları bulunmaktadır. Her yeni veri noktası için tüm veri setini taraması nedeniyle, özellikle büyük veri setlerinde hesaplama maliyeti yüksek olabilmekte ve algoritmanın yavaş çalışmasına neden olabilmektedir [69]. Ayrıca, büyük veri kümelerinde ve dengesiz sınıflarda performans düşüşü yaşanabilmektedir. Bu nedenle, özellikle çok sayıda örneği olan veri setlerinde başka yöntemlerle kombine edilmesi önerilmektedir [70].

## 4.2 Prenatal Tanıda Kullanılan Makine Öğrenmesi ve Yapay Zeka Uygulamaları

Prenatal tanı alanında son yıllarda yaşanan teknolojik gelişmeler, özellikle yapay zeka (AI) ve makine öğrenmesi (ML) algoritmalarının entegrasyonu ile yeni bir boyut kazanmıştır. NIPT yaygınlaşması ve cffDNA analizlerindeki ilerlemeler, risk değerlendirme ve tanı süreçlerini önemli ölçüde değiştirmiştir [71–74]

Prenatal tanı risk değerlendirmesinde, yapay zeka destekli sistemler çoklu biyobelirteç analizlerini entegre ederek daha hassas tahminler yapabilmektedir. Bu sistemler özellikle erken doğum riskinin belirlenmesinde önemli başarı göstermiştir. Geleneksel yöntemlerde kullanılan tek parametreli değerlendirmelerin ötesine geçilerek, metabolomik, proteomik ve sitokin analizleri gibi çoklu omik verilerinin makine öğrenmesi algoritmaları ile değerlendirilmesi, risk tahmininde daha yüksek doğruluk sağlamıştır [75].

Risk değerlendirmesinde kullanılan başlıca biyobelirteçler şunlardır: gebelikle ilişkili plazma protein-A (PAPP-A) konsantrasyonu, beta insan koryonik gonadotropini ($\beta$-hCG), dimerik inhibin A ve alfa-fetoprotein (AFP) seviyeleridir [76].

Bu biyobelirteçlerin özellikle transvajinal ultrasonografi bulguları ile birlikte değerlendirilmesi, risk tahmininin güvenilirliğini artırmıştır. Serviks uzunluğu 15 mm'den kısa olan gebelerde, biyobelirteç verileri ile entegre edilen demografik veriler (maternal yaş, boy, pre-gebelik vücut kitle indeksi, gelir düzeyi, eğitim durumu, ırk ve göçmenlik durumu) ve klinik veriler (önceki düşük sayısı, sigara kullanımı, alkol tüketimi, folik asit kullanımı, ilk trimester ziyareti, doğum öncesi sağlık hizmeti sağlayıcı tipi, mevcut sağlık durumları, diyabet ve mevcut ruhsal sağlık durumları) makine öğrenmesi algoritmalarıyla analiz edilmiştir [76]. Bu entegre analiz yaklaşımı, geleneksel yöntemlere kıyasla daha yüksek öngörü değeri sağlamıştır [75].

Bu verilere ek olarak, gebeliğe özgü klinik parametreler arasında gebelik oluşum tipi, ilk trimesterdeki gestasyonel kilo alımı ve gebelik komplikasyonları da risk değerlendirmesinde önemli faktörler olarak belirlenmiştir [76]. Makine öğrenmesi algoritmaları, bu çok sayıda değişkeni eş zamanlı analiz ederek, erken doğum riskinin daha hassas ve doğru tahmin edilmesine olanak sağlamıştır

Fetal görüntüleme verilerinin analizi için geliştirilen derin öğrenme algoritmaları, özellikle plasental değerlendirmede önemli ilerlemeler sağlamıştır. İlk trimesterde plasental hacmin otomatik ölçümü için geliştirilen derin öğrenme tabanlı üç boyutlu ultrason segmentasyon tekniği, zamanında doğan bebeklerin doğum ağırlıklarıyla güçlü korelasyon göstermiştir. Bu teknoloji, potansiyel büyüme kısıtlılığı olan fetüslerin erken dönemde belirlenmesine olanak sağlamaktadır [75].

Genomik alanda yapay zeka uygulamaları, özellikle DNA dizileme ve fenotipleme süreçlerini hızlandırmıştır. Hızlı otomatize DNA dizileme sistemleri ve derin fenotipleme ile elektronik sağlık kayıtlarının yapay zeka destekli analizi, dizileme sonuçlarının yorumlanma süresini önemli ölçüde kısaltmıştır. Yenidoğan yoğun bakım ünitelerinde tanı süresinin ortalama 20 saate kadar düşürülmesi, bu teknolojinin klinik uygulamadaki etkinliğini göstermektedir [75].

Embriyo seçiminde kullanılan derin öğrenme algoritmaları, *time-lapse* video görüntüleme teknolojisi ile elde edilen verilerin analizi sonucunda %96.4'lük bir doğruluk oranına ulaşmıştır. Bu sistemler, embriyo kalite değerlendirmesinin yanı sıra, kültür koşullarının optimizasyonu ve laboratuvar verimliliğinin artırılmasına da katkı sağlamaktadır [75].

Yapay zeka ve makine öğrenmesi teknolojilerinin prenatal tanıda kullanımı, risk değerlendirme ve tanı süreçlerini daha hassas ve kişiselleştirilmiş hale getirmektedir. Çoklu biyobelirteç analizleri ve görüntüleme teknikleriyle entegre edilen bu sistemler, minimal invaziv yaklaşımların geliştirilmesine ve daha etkin tedavi stratejilerinin belirlenmesine olanak sağlamaktadır [75]. Gelecekte, bu teknolojilerin daha da gelişmesiyle, prenatal tanı ve tedavi yaklaşımlarının doğruluğunun ve etkinliğinin artması beklenmektedir. Ve bu teknolojilerin klinik pratiğe başarılı bir şekilde entegre edilebilmesi için, prospektif çalışmalar yapılması, standartların belirlenmesi ve insan-makine etkileşiminin optimize edilmesi gibi konularda kapsamlı araştırmalara ihtiyaç duyulmaktadır.

# 5
# MATERYAL-METOT

## 5.1 Veri Kaynakları ve Toplanma Süreci

Çalışmada kullanılan veriler, Başakşehir Çam ve Sakura Şehir Hastanesi Genetik Hastalıkları Değerlendirme Merkezi'nde 2022-2024 yılları arasında prenatal tanı süreçlerinden elde edilen QF-PCR sonuçlarını kapsamaktadır. Araştırma protokolü, İstanbul Başakşehir Çam ve Sakura Şehir Hastanesi Klinik Araştırmalar Etik Kurulu tarafından 05.03.2025 tarihinde KAEK/2025-78 protokol kodu ile onaylanmış ve tüm veriler anonimleştirilerek çalışmaya dahil edilmiştir. Çalışma popülasyonu, son iki yıl içinde QF-PCR testi yaptıran 354 prenatal hasta örneğinden oluşmaktadır. Hasta seçiminde eksik parametreli vakalar çalışma dışı bırakılmıştır. Çalışma popülasyonu; yaş, boy, kilo gibi demografik veriler; ultrasonografi bulguları ve biyokimyasal tarama test sonuçları gibi klinik veriler ile QF-PCR sonuçlarından oluşan genetik veriler olmak üzere üç ana kategoride sınıflandırılmıştır.

### 5.1.1 Çalışma Popülasyonunun Sınıflandırılması

Verilerin güvenilir ve doğru bir analiz sağlayabilmesi için veri toplanması ve düzenlenmesi yapılmıştır. Çalışmada veri bütünlüğünü sağlamak amacıyla korumacı bir yaklaşım benimsenmiş olup, başlangıçta toplanan 412 vakanın kapsamlı bir değerlendirmesi yapılmıştır. Bu süreçte, hasta anamnez formlarına dayalı ve Hastane Bilgi Yönetim Sistemi (HBYS) üzerinden her bir hastaya ait demografik ve klinik veriler, hasta protokol numarası baz alınarak manuel bir şekilde Excel veri dosyasına aktarılmıştır. Makine öğrenmesi analizlerinde yalnızca demografik ve klinik bilgiler ve genetik sonuçlar kullanılmış, hastaların kişisel tanımlayıcı bilgileri analizlere dahil edilmemiştir. Bu şekilde, hasta mahremiyeti korunarak veriler tamamen anonimleştirilmiş halde değerlendirilmiştir.

Bu süreç hastaların başvuru anından gebelik sonuçlanmasına kadar olan tüm klinik kayıtları, laboratuvar sonuçları, tedavi süreçleri; Tablo 5.1'de belirlediğimiz parametreler çerçevesinde işleme alınmıştır.

**Tablo 5.1** Çalışmada kullanılacak parametreler ve açıklamaları

| **Parametre** | **Açıklama** |
|---|---|
| Yaş | Hastanın yaşı,trizomi riskini etkileyebilecek önemli bir demografik bilgidir. |
| Cinsiyet | Hastanın cinsiyeti |
| Akrabalık | Hastanın eş/partner ile akrabalık durumu (var/yok), genetik hastalık risklerini artırabilir. |
| Gebelik Türü | Gebeliğin türü (Normal, tüp bebek veya donasyon, yumurta/sperm bağışı). |
| Gebelik Sayısı | Hastanın toplam gebelik sayısı (canlı ve düşük dahil). Annenin doğurganlık geçmişini değerlendirir. |
| Doğum Sayısı | Hastanın gerçekleştirdiği toplam canlı doğum sayısı |
| Düşük Sayısı | Hastanın daha önce yaşadığı düşüklerin sayısı |
| Çocuk Sayısı | Hastanın mevcut yaşayan çocuk sayısı |
| Bilinen Hastalık | Hastanın sahip olduğu kronik veya geçmişte tanısı konulmuş bilinen hastalıklar |
| Geçirilen Operasyon | Hastanın daha önce geçirdiği cerrahi işlemler, özellikle üreme organlarına yönelik operasyonlar önemlidir. |
| Kullanılan İlaç | Hastanın kullandığı reçeteli veya reçetesiz ilaçlar. İlaçların fetüs üzerindeki etkileri değerlendirilir. |
| Alerji | Hastanın sahip olduğu alerjik reaksiyonlar |
| Sigara | Hastanın sigara kullanımı durumu (aktif veya geçmiş kullanım). Sigara gebelikte birçok komplikasyona neden olabilir. |
| Gebelik Haftası | Mevcut gebelik haftası (örneğin 12. hafta) |
| NF | Ense kalınlığı ölçümü, genetik anomali taramalarında kullanılan önemli bir ultrasonografi parametresidir. |
| USG | Ultrasonografi sonuçları. Bebeğin anatomik ve fizyolojik durumunun değerlendirilmesinde kullanılır. |
| Tarama Testi | Yapılan tarama testi varlığı. Bu sonuçlar genetik risk değerlendirmesi için kullanılır. |
| Tarama Testi Türü | Gebelikte uygulanan tarama testi türü (örneğin ikili test, üçlü test vb.). Bu bilgi, test sonuçlarının değerlendirilmesi için önemlidir. |
| Tarama Testi Riski | Tarama testlerinin sonucunda hesaplanan risk oranı (örneğin Düşük-orta-yüksek riski). Risk değerlendirmesi ve ileri tanı testlerinin gerekliliği bu bilgiye dayanır. |
| QF-PCR Sonuç | QF-PCR testi ile elde edilen sonuçlar. Bu sonuçlar genetik anomalilerin varlığına ilişkin bilgi verir. |
| Numune Tipi | Tanı için kullanılan biyolojik materyal türü (örneğin amniyon sıvısı, koryon villus örneği). |
| Gebelik Sonucu | Gebeliğin sonucu (örneğin normal doğum, düşük, terminasyon, sezaryen). Gebelik yönetimi ve sonuçların değerlendirilmesi için önemlidir. |

Belirlenen parametrelerde karşılığı olmayan, eksik veya yetersiz olan vakalara, modelin performansını olumsuz etkilemesini önlemek amacıyla vaka dışlama

yöntemi uygulanmıştır. Bu kapsamda başlangıçtaki 412 vakanın 58'i çalışma dışı bırakılmış ve analizler 354 vaka üzerinden yürütülmüştür. Tam olmayan hastalar veri setinin küçülmesine rağmen, bu yaklaşım analiz sonuçlarının güvenilirliğini ve klinik uygulanabilirliğini önemli ölçüde artırmıştır.

Bu parametreler, gebelik sürecinde hastanın genel durumunu, tıbbi öyküsünü ve genetik değerlendirmesini kapsamlı bir şekilde ele almak amacıyla üç ana kategoriye ayrılmıştır. Her kategori, gebelik yönetimi ve sonuçlarının optimize edilmesi için özel bir bakış açısı sunar.

#### 5.1.1.1 Demografik Veriler

İlk kategoride demografik veriler yer almakta olup, hastaların yaşı, cinsiyeti, gebelik sayısı, doğum sayısı, çocuk sayısı, akrabalık durumu, gebelik türü ve sigara kullanımı gibi temel bilgiler kaydedilmiştir. Bu veriler, hastaların genel sağlık durumunu ve potansiyel risk faktörlerini değerlendirmek amacıyla sistematik olarak toplanmıştır.

#### 5.1.1.2 Klinik Veriler

İkinci kategoride klinik veriler bulunmaktadır. Bu kapsamda ultrason bulguları, biyokimyasal tarama testleri (ikili ve üçlü testler), ebeveynlerin bilinen genetik hastalık geçmişleri, geçirilmiş operasyonlar, kullanılan ilaçlar, alerjik reaksiyon öyküsü, gebelik haftası, düşük sayısı, tarama testi ile belirlenen risk düzeyi ve gebelik sonlanma durumu gibi detaylı klinik bilgiler bir araya getirilmiştir. Bu veriler, genetik risk faktörlerinin yanı sıra klinik olarak belirlenmiş risklerin bütüncül değerlendirilmesi amacıyla kullanılmıştır.

#### 5.1.1.3 Genetik Veriler

Üçüncü kategori genetik verileri içermektedir. QF-PCR sonuçları ve numune türleri bu kategorinin temelini oluşturmaktadır. Her QF-PCR testi, 21, 18, 13, X ve Y kromozomları üzerinde anöploidi tespiti yapılarak değerlendirilmiştir. Çalışmada kullanılan numune türleri arasında amniyon sıvısı, abort materyali, kordon kanı, fetal kan, cilt biyopsisi, kordosentez, periferik kan, kardiyak kan ve koryon villus biyopsisi (CVS) yer almıştır. Ek olarak, karyotipleme ve array-CGH (Mikrodizi Karşılaştırmalı Genomik Hibridizasyon) sonuçları da toplanarak, genetik doğrulayıcı bilgiler olarak ve istatistiksel karşılaştırmalarda paralel değerlendirme yapmak için kullanılmıştır. QF-PCR sonuçlarına istinaden çalışmadaki anöploidi çıkan hastaların maternal ya da paternal sınıflandırmaları da yapılmıştır.

## 5.2 Veri Yönetimi ve Ön İşleme

### 5.2.1 Veri Doğrulama Süreci ve Kullanılan R Kodu

Veri doğrulama sürecinde, tamamlanmayan verilerin Excel hücresinde kontrol edilmesi amacıyla Şekil 5.1'de gösterilen R kodu kullanılmıştır. Bu kod, veri setindeki her bir kolonda tamamlanmamış değerlerin olup olmadığını belirlemek, hatalı girişleri tespit etmek ve veri tutarsızlıklarını değerlendirmek için tasarlanmıştır.

```r
library(readxl)
library(dplyr)

# Veri Seti Yükleme-Okuma
df <- read_excel("C:/Users/ASUS/Desktop/data_use.xlsx")
# Veriyi gözden geçirme
str(df)
# NA olan satırların tespiti
na_rows <- df[is.na(df$nf), ]
# Eksik değerlerin kontrolü
missing_summary <- sapply(df, function(x) sum(is.na(x)))

# Değer aralıklarının kontrolü
age_check <- df$yas >= 16 & df$yas <= 48
pregnancy_week_check <- df$gebelik_haftasi >= 0 &
    df$gebelik_haftasi <= 40
pregnancy_number_check <- df$gebelik_sayisi >= 0 &
    df$gebelik_sayisi <= 15

# Tutarsızlık kontrolü
consistency_check <- df$dogum_sayisi <= df$gebelik_sayisi

# Tutarlılık kontrolü
df <- df %>%
    mutate(Gebelik_Dogum_Hata =
    ifelse(dogum_sayisi > gebelik_sayisi, "Hatalı", "Geçerli"))
```

**Şekil 5.1** Veri setinde eksik değerlerin tespiti, değer aralıklarının kontrolü ve tutarsızlıkların kontrolü için kullanılan R kodu

#### 5.2.1.1 Veri Doğrulama Kriterleri

İlk aşamada, değişkenlerin klinik ve mantıksal tutarlılığı kontrol edilmiştir.

- Demografik Tutarlılık:

  - Anne yaşı değerlerinin klinik prenatal tanı yaş aralığında olması (veri setimizdeki aralık: 16-48 yaş)
  - Gebelik haftasının normal gebelik süreci içinde olması (6-40 hafta arası).
- Gebelik İle İlgili Parametreler:
  - Gebelik sayısının mantıksal üst sınırı içinde olması,
  - Doğum sayısının gebelik sayısından fazla olmaması,
  - Düşük sayısının gebelik sayısı ile tutarlı olması.
- Klinik Bulgular:
  - USG bulgularının standart sınıflandırma sistemine uygunluğu,
  - Tarama test sonuçlarının geçerli risk değerleri içinde olması,
  - QF-PCR sonuçlarının kabul edilen tanı kategorilerinde olması.

### 5.2.2 Veri Temizleme ve Standardizasyon

Veri temizleme ve standardizasyon süreçleri, makine öğrenmesi algoritmaları ve istatistiksel analizlerde doğru sonuçlar elde edilmesi için Excel ile gerçekleştirilmiştir.

#### 5.2.2.1 Kategorik Verilerin Standardizasyonu

Kategorik verilerdeki yazım farklılıkları, veri gruplarının doğru şekilde sınıflandırılması, tekrarlayan kategorilerin birleştirilmesi ve analizlerin güvenilirliğinin artırılması amacıyla standart bir forma dönüştürülmüştür. Kategorik verilerdeki yazım farklılıkları, analizlerde tutarlılık sağlamak amacıyla şu şekilde düzenlenmiştir:

- Gebelik Sonucu: "Normal doğum", "ND", "normal" gibi farklı yazımlar "Normal Doğum" olarak standardize edilmiştir.
- Numune Tipi: "AS", "amniyon", "Amnion sıvısı" gibi yazımlar "Amniyosentez" olarak düzenlenmiştir.
- USG Bulguları: "Normal", "N", "normal" gibi yazımlar tek bir formatta standardize edilmiştir.

Kategorik verilerdeki yazım farklılıkları önce metin olarak standardize edilmiş, ardından algoritmalarında kullanılmak üzere sayısal kodlamaya dönüştürülmüştür:

- Gebelik sonucu kategorisi standardize edildikten sonra beş sınıfa ayrılmıştır:
  - CS (Sezaryen Doğum), (0)
  - Normal Doğum (1)
  - NA (Mevcut değil) (2)
  - Terminasyon (3)
  - Abort (4)
- Numune tipi kategorisi beş ana grupta kodlanmıştır:
  - Amniyosentez (0)
  - Abort materyali (1)
  - Koryon villus biyopsisi (2)
  - Kordon kanı (3)
  - Diğer numune türleri (4)
- USG bulguları, veri standardizasyonundan sonra klinik sınıflandırmaya uygun şekilde kodlanmıştır:
  - Normal (1)
  - Kistik lezyonlar (2)
  - Yüz anomalileri (3)
  - Kalp ve damar anomalileri (4)
  - Renal anomaliler (5)
  - Merkezi sinir sistemi anomalileri (6)
  - İskelet sistemi anomalileri (7)
  - Diğer sistem anomalileri (8)
  - Gastrointestinal anomaliler (9)
  - Çoklu anomaliler (10)
  - Bilinmeyen (11)
- Tarama testi ile belirlenen risk düzeyi, biyokimyasal parametreler (örneğin, serbest $\beta$-hCG, PAPP-A, AFP) ve ultrason bulgularının (ense kalınlığı ve baş-popo mesafesi gibi) analiz edilmesiyle hesaplanmıştır. Anne yaşı ve klinik geçmiş gibi ek faktörler de değerlendirmeye dahil edilerek kombine bir risk skoru aşağıdaki yöntemle hesaplanmıştır:

*Kombine Risk Skoru = (Yaşa bağlı temel risk) * ($\beta$-hCG faktörü) * (PAPP-A faktörü) * (Ultrason faktörü)*, [77, 78].

Elde edilen risk skorları, en yüksek risk grubundan en düşük risk grubuna doğru sıralanarak beş eşit aralığa bölünmüş ve 'düşük', 'orta-düşük', 'orta', 'yüksek' ve 'çok yüksek' kategorilerine atanmıştır [79]. Bu kategoriler sırasıyla 1, 2, 3, 4 ve 5 olarak sayısal kodlamaya dönüştürülmüştür. Bu yöntem, risk seviyelerinin daha objektif bir şekilde sınıflandırılmasını sağlayarak analiz süreçlerinde standart ve nesnel bir değerlendirme sunmayı amaçlamaktadır.

Kategorik değişkenlerin sayısal kodlama ile dönüştürülmesi, istatistiksel analizler ve makine öğrenmesi modellerinde veri ön işlemenin temel aşamalarından birini oluşturmaktadır. Bu dönüşüm sürecinde, kategorik değişkenler sıralı sayısal değerlere (örneğin, 1, 2, 3 gibi) dönüştürülerek analiz için uygun hale getirilmektedir.

Label encoding yöntemiyle gerçekleştirilen bu işlemde, her kategoriye belirli bir sayısal değer atanır. Örneğin, gebelik sonucu değişkeni için "normal doğum" kategorisi 1, "sezaryen" kategorisi 2 ve "terminasyon" kategorisi 3 olarak kodlanabilir.

Bu bağlamda, R programlama dilinin sunduğu faktör veri yapısı, kategorik değişkenlerin bu şekilde işlenmesini kolaylaştırmakta ve istatistiksel analizlerde doğru sonuçlar elde edilmesine olanak tanımaktadır. Faktör yapısı, R'ın yerel kategorik veri tipi olarak, hem veri tipi doğruluğunu hem de model uyumluluğunu sağlamaktadır.

#### 5.2.2.2 Excel Araçları ile Veri Temizleme

Excel'in çeşitli araçları, veri temizleme sürecinde etkin bir şekilde kullanılmıştır:

- Veri Doğrulama (Data Validation) ile kategorik değişkenler için açılır listeler oluşturulmuştur.
- Koşullu Biçimlendirme ile aykırı değerler tespit edilmiştir.
- TRIM fonksiyonu ile hücrelerin başındaki ve sonundaki gereksiz boşluklar temizlenmiştir (örneğin: " Normal " -> "Normal").
- Pivot tablolar ile veri tutarlılığı, kategorik değişkenlerdeki benzersiz değerler kontrol edilmiştir.

Bu veri temizleme süreci, Excel'in gelişmiş özelliklerinden yararlanılarak sistematik ve tekrar edilebilir bir şekilde gerçekleştirilmiştir. Özellikle kısayolların (bul ve değiştir: Ctrl + H, filtre özelliği: Ctrl + Shift + L gibi) ve yerleşik fonksiyonların etkin kullanımı, sürecin verimliliğini artırmıştır.

#### 5.2.2.3 Sayısal Veri Düzenlemeleri

Sayısal veriler aşağıdaki yöntemlerle düzenlenmiştir:

- Gebelik haftası: Gün olarak girilen değerler haftaya çevrilmiştir. ROUNDDOWN fonksiyonu kullanılmıştır.
  - Formül: =ROUNDDOWN(gün/7, 0). Buradaki 0 değeri, ondalık basamak sayısını belirtir.
  - Örnek: 245 gün $\rightarrow$ 35 hafta
- Yaş hesaplaması: Yaşlar, doğum ve işlem tarihi kullanılarak hesaplanmıştır. DATEDIF fonksiyonu kullanılmıştır.
  - Formül: =DATEDIF(doğum-tarihi, işlem-tarihi, "y"), "y" parametresi yıl cinsinden sonuç almak için kullanılmıştır.
  - Örnek: =DATEDIF("1990-01-01", "2023-12-31", "y") $\rightarrow$ 33 yaş
- Risk değerleri: 1/x formatında standardize edilmiştir.

#### 5.2.2.4 Aşırı Uç Değerlerin Tespiti ve İşlenmesi

Veri setindeki aşırı uç değerlerin tespiti ve işlenmesi, analizlerin güvenilirliği açısından kritik bir aşama olarak değerlendirilmiştir. Veri setindeki sıra dışı değerler, hem istatistiksel analiz hem de klinik değerlendirme ile tespit edilmiştir:

- Anne Yaşı: 16-48 yaş aralığına uymayan değerler sıra dışı olarak değerlendirilmiştir.
- Gebelik Haftası: 6-40 hafta aralığına uymayan değerler analiz dışı bırakılmıştır.
- Risk Değerleri: Klinik sınırlar dışında kalan risk değerleri, yeniden gözden geçirilmiştir.

Veri setindeki sıra dışı görünen değerler, uzman görüşü ve hasta dosyalarının hastane bilgi sisteminden doğrulanması ile değerlendirilmiştir. Klinik olarak mümkün olan

ancak nadir görülen durumlar (örneğin, 45 yaşındaki gebe birey) analizde korunurken, istatistiksel tutarsızlık gösteren değerler veri setinden çıkarılmıştır.

#### 5.2.2.5 Verilerin Normalizasyonu ve Standardizasyonu

Farklı ölçeklerdeki değişkenlerin bir arada kullanıldığı analizlerde karşılaştırılabilirlik sorunlarının önüne geçmek ve makine öğrenmesi algoritmalarının optimum performans gösterebilmesi için, veri standardizasyonu ve normalizasyonu işlemleri üç farklı yöntemle gerçekleştirilmiştir: sürekli değişkenler için z-skor dönüşümü, belirli aralıktaki veriler için min-max normalizasyonu ve büyük sayısal değerler için ondalık ölçekleme uygulanmıştır.

- **Standartlaştırma (Z-skor dönüşümü):** Sürekli değişkenlerin farklı ölçek yapılarından kaynaklanan karşılaştırılabilirlik sorunlarını gidermek amacıyla z-skor standartlaştırma yöntemi uygulanmıştır. Bu kapsamda yaş, gebelik haftası ve doğum sayısı gibi sürekli değişkenler, kendi ortalamalarından çıkarılıp standart sapmalarına bölünerek dönüştürülmüştür. R programlama dilinin `scale()` fonksiyonu kullanılarak gerçekleştirilen bu işlem sonucunda, her bir değişkenin ortalaması 0, standart sapması 1 olacak şekilde yeniden ölçeklendirilmiştir. Bu dönüşüm, değişkenlerin orijinal dağılım yapısını korurken ölçek farklılıklarını ortadan kaldırmış, böylece makine öğrenmesi algoritmalarının veri setindeki örüntüleri daha etkili bir şekilde tespit edebilmesine olanak sağlamıştır.

- **Ondalık Ölçekleme:** Büyük sayılar, daha kolay anlaşılabilir hale getirilmek için belirli bir oranla küçültülmüştür. Örneğin, 1000 gibi büyük bir değer, 10'a bölünerek daha küçük bir formatta sunulmuştur. Bu da orjinal verinin yorumlanabilirliğini korur.

- **Min-Max Normalizasyonu:** Risk değerleri gibi belirli bir aralıkta olması gereken değişkenler için min-max normalizasyonu uygulanmıştır. Bu yöntemde her bir veri noktası, veri setindeki en küçük ve en büyük değerler referans alınarak 0 ile 1 arasında yeniden ölçeklendirilmiştir. Örneğin, anne yaşı değişkeni için veri setindeki en genç (16 yaş) ve en yaşlı (48 yaş) anneler referans noktaları olarak kullanılmış, tüm yaş değerleri bu aralıkta normalize edilmiştir. Bu dönüşüm sonucunda 25 yaşındaki bir anne, 0-1 ölçeğinde yaklaşık 0.28 değeri ile temsil edilmiştir. Bu normalizasyon yaklaşımı, farklı ölçeklerdeki değişkenlerin aynı aralıkta ifade edilmesini sağlayarak, makine öğrenmesi modellerinde değişkenler arası karşılaştırılabilirliği ve tutarlılığı artırmıştır.

Uygulanan standardizasyon ve normalizasyon işlemleri, çalışmanın analiz sürecinde çeşitli açılardan kritik rol oynamıştır. Bu dönüşümler, özellikle lojistik regresyon gibi mesafe ve ağırlık temelli algoritmaların performans optimizasyonunda önemli katkı sağlamıştır. Farklı ölçeklerdeki değişkenlerin karşılaştırılabilir hale getirilmesi, model parametrelerinin daha etkin öğrenilmesini sağlamış ve lojistik regresyonda kullanılan optimizasyon süreçlerinin daha hızlı ve kararlı bir şekilde sonuçlanmasına olanak tanımıştır. Bu sayede, modellerin tahmin doğruluğu artırılmış ve hesaplama verimliliği optimize edilmiştir.

## 5.3 Verilerin Sistematik Sınıflandırması

Çalışmada verilerin sistematik sınıflandırılması, kompleks kromozomal varyasyonların, ultrason bulgularının, QF-PCR sonuçlarının, gebelik sonucunun, geçirilen operasyonların anlaşılabilir kategorilere dönüştürülmesinde kritik bir rol oynamaktadır. Bu bağlamda, karyotip mutasyonlarının yedi ana kategoride (normal, monozomi, trizomiler, diğer anomaliler, translokasyonlar, insersiyonlar ve inversiyonlar) standardize edilmesi Tablo 5.2'de görülmektedir. Hem klinik değerlendirme süreçlerini optimize etmekte hem de veri analizi için yapısal bir temel oluşturmaktadır. Her bir kategorinin numerik kodlarla eşleştirilmesi ve alt grupların sistematik organizasyonu, genetik anomalilerin örüntü analizini kolaylaştırırken, makine öğrenmesi algoritmalarının eğitimi için de yapılandırılmış bir veri seti sunmaktadır.

USG verilerinin sınıflandırmaları ve sistematik şekilde gruplandırılması, fetüsün anatomik ve fizyolojik anomalilerinin değerlendirilmesi sürecinde önemli bir yapı taşını oluşturmaktadır. Ultrasonografi bulguları, kistik lezyonlar, yüz anomalileri, kalp ve damar anomalileri gibi kategorilere ayrılmış olup, her bir kategori içinde spesifik bulgular ve ilişkili ek semptomlar tanımlanmıştır (Tablo 5.3). Bu sınıflandırma, hem klinik karar süreçlerini optimize etmekte hem de tanısal doğruluğu artırmak amacıyla detaylı bir temel sunmaktadır. Her bir USG bulgusunun kategorilere göre kodlanması ve alt grupların organize edilmesi, ultrasonografi sonuçlarının daha tutarlı bir şekilde yorumlanmasına olanak tanırken, verinin düzenli bir biçimde analiz edilmesini sağlamaktadır. Aynı zamanda, makine öğrenmesi algoritmalarının eğitiminde kullanılacak yapılandırılmış bir veri seti oluşturularak, bu algoritmaların prenatal tanıda kullanılmak üzere daha etkin modeller geliştirmesine olanak tanımaktadır. Bu yaklaşım, USG verilerinin sınıflandırılması ve ilgili ek semptomların düzenli bir veri yapısına dönüştürülmesi yoluyla, genetik ve klinik analizlerin entegrasyonunu kolaylaştırmaktadır.

**Tablo 5.2** Kompleks kromozamal varyasyonların kategorize edilmesi

| Kategori Numarası | Karyotipleme Sonucu | Kategorik İsimlendirme |
|---|---|---|
| 1 | Normal | Normal |
| 2 | 45,X | Monozomi |
| | mos 45,X[44]/46,XX[6] | |
| 3 | 47,–,+13 | Trizomiler |
| | 47,–,+18 | |
| | 47,–,+21 | |
| | 47,–,t(6;22)(q25.3;q11.2),+21 | |
| 4 | 46,XYY | Diğer Anomaliler |
| | 47,XXY | |
| | 69,— | |
| 5 | 46,–,(13;14)(q10;q10),+13 | Translokasyonlar |
| | 46,–,(14;21)(q10;q10),+21 | |
| | 46,–,(6)(q21q22.3) | |
| | 46,XX,t(11;22)(q24;q12) | |
| | 46,XY,+13,der(13;13)(q10;q10) | |
| | 46,XX,der(13;14)(q10;q10),+13 | |
| 6 | 46,–,ins(9)(p11q13) | İnsersiyonlar |
| 7 | 46,XX,inv(12)(p11.2q13) | İnversiyonlar |

Kategorik değişkenler, verinin yapısına uygun olarak kodlanmıştır. Tüm kategorik değişkenler, analizlerde doğru şekilde işlenebilmesi için R programlama dilinde `as.factor()` fonksiyonu kullanılarak kategorik veri tipine dönüştürülmüştür. Çalışmamızda makine öğrenmesi algoritmalarının hedef değişkeni olan QF-PCR test sonuçları, ikili (binary) sınıflandırma problemi olarak ele alınmıştır. QF-PCR test sonuçları, normal olan örnekler '0' ve kromozomal anomali (anöploidi) tespit edilen örnekler '1' olacak şekilde kodlanmıştır. Bu sınıflandırma, makine öğrenmesi algoritmalarının prenatal taramada kromozomal anomalileri tahmin etme performansını değerlendirmek için temel oluşturmuştur.

## 5.4 Çalışma Ortamı ve Kullanılan R Paketleri

Donanım ve İşletim Sistemi: Çalışmada Windows 11 işletim sistemine sahip, 64-bit mimaride ve 32 GB RAM kapasiteli bir bilgisayar kullanılmıştır.

Yazılım Ortamları: Analizlerin kapsamlı ve etkin bir şekilde gerçekleştirilebilmesi için R 4.3.3 sürümü tercih edilmiştir. R programlama dili, analiz süreçlerinde istatistiksel modelleme, makine öğrenmesi ve görselleştirme için kullanılmıştır. Çalışmada, analizlerin etkin bir şekilde yürütülmesi ve verilerin kapsamlı bir şekilde incelenebilmesi için çeşitli R paketlerinden yararlanılmıştır.

**Tablo 5.3** USG sınıflandırma tablosu

| Kategori Numarası | Kategorik İsimlendirme | USG Sınıflandırmaları |
|---|---|---|
| 1 | Normal | - |
| 2 | Kistik Higroma ve Diğer Kistik Lezyonlar | Kistik higroma<br>Bilateral/Unilateral koroid pleksus kistleri<br>Mega sisterna magna-Polihidramnios-Oligohidramnios<br>Septal kistik higroma |
| 3 | Nazal Kemik ve Yüz Anomalileri | Nazal kemik hipoplazisi<br>Yarık dudak damak<br>Mikroftalmia<br>Tek burun deliği<br>Bilateral Proptozis/Ekzoftalmus<br>Düşük Kulak |
| 4 | Kalp ve Damar Anomalileri | Arteriozus<br>ASD-VSD-ARSA-AVSD<br>Pulmoner arter darlığı<br>Hiperekojen kardiyak odak<br>Trikuspid regürjitasyon<br>Fallot Tetralojisi |
| 5 | Renal ve Üriner Sistem Anomalileri | Bilateral pelviektazi<br>Renal pelvis genişliği<br>Multikistik displastik böbrek<br>Konjenital adrenal hiperplazi |
| 6 | Serebral ve Merkezi Sinir Sistemi Anomalileri | Ventrikulomegali<br>Holoprosensefali<br>Spina bifida<br>Dandy-Walker malformasyonu-Vermian Agenezi<br>Korpus Kallosum ve CSP Görülmesi<br>NTD (Nöral Tüp Defekti) |
| 7 | İskelet Anomalileri | İskelet displazisi<br>Kısa FL (Femur Length)<br>Pes ekinovarus<br>Omurga Deformiteleri-Sandal Gap-Rocker bottom |
| 8 | Diğer Sistem Anomalileri | Ductus Venosus Anomalileri(Ters A Dalgası)<br>Çift kabarcık bulgusu<br>Hipertelorizm<br>Timüs Hipoplazik<br>Fetal Solunum ve Ekstremite Hareketleri |
| 9 | Gastrointestinal Sistem Anomalileri | Bağırsaklar Hiperekojen<br>Umbilikal Kord Kesesi Hernisi<br>Mide, Safra Kesesi Yokluğu<br>Diyafram Hernisi |
| 10 | Çoklu Anomaliler | Birden fazla sistemde anomali olması |
| 11 | Bilinmeyen | - |

Çalışmamızda kullanılan başlıca R paketleri ve işlevleri şunlardır:

- `readxl`: paketi, R programlama ortamında Excel dosyalarını (xlsx ve xls formatları) verimli bir şekilde içe aktarmak ve `read_excel()` fonksiyonu aracılığıyla veri analizi süreçlerini optimize etmek için kullanılmıştır.

- `caret`: Makine öğrenmesi modellerinin geliştirilmesi, optimize edilmesi ve değerlendirilmesi için kapsamlı bir araç seti sunmaktadır. Proje kapsamında, özellikle `createDataPartition` fonksiyonu kullanılarak eğitim ve test veri setlerinin oluşturulması sağlanmıştır.

- `randomForest`: Karar ağaçları yöntemini temel alan random forest algoritmasını uygular ve sınıflandırma ile regresyon analizlerinde kullanılmıştır.

- `rpart`: Karar ağaçları oluşturmak için kullanılan bir diğer pakettir; sınıflandırma ve regresyon ağaçları için hızlı bir yöntem sağladı.

- `rpart.plot`: `rpart` tarafından oluşturulan karar ağaçları, daha anlaşılır bir şekilde görselleştirilmiştir.

- `ggplot2`: Verilerin görselleştirilmesi için kapsamlı bir grafik ve çizim sistemi sunmaktadır. Bunun yanı sıra, bu proje kapsamında kullanılan diğer kütüphanelerle birlikte veri analizi süreçlerini desteklemektedir.

- `nnet`: Sinir ağları ve çok sınıflı lojistik regresyon(multinom modeli) uygulamaları için kullanılmıştır.

- `pROC`: ROC eğrileri ve AUC(Alan Altında Eğri) değerlerinin hesaplanması ve görselleştirilmesi için kullanılmıştır.

- `ROCR`: Sınıflandırma algoritmalarının performansını değerlendirmek için ROC eğrileri ve diğer ölçümleri görselleştirmek amacıyla kullanılmıştır.

## 5.5 İstatistiksel Analizler

Bu retrospektif çalışmada, prenatal tanı ve tarama testlerinin gebelik sonuçları üzerindeki etkisi istatistiksel yöntemlerle değerlendirilmiştir. Çalışmanın temel amacı, prenatal tarama ve tanı testlerinin gebelik sonuçları ile ilişkisini incelemek ve bu ilişkinin klinik açıdan anlamlı olup olmadığını belirlemektir.

Çalışmaya dahil edilen gebelerin yaş, gebelik haftası ve tarama testi kullanımı gibi demografik ve klinik özellikleri değerlendirilmiştir. Bu değişkenler, ortalama, medyan, standart sapma (SS), minimum ve maksimum değerler kullanılarak tanımlayıcı istatistiklerle özetlenmiştir (Tablo 7.4).

Çalışmada, farklı yaş gruplarında tarama testlerinin uygulanma oranları ve bu testlerin gebelik sonuçlarına etkileri istatistiksel olarak analiz edilmiştir. Yaş grupları ile tarama testi varlığı arasındaki ilişki, Ki-kare testi kullanılarak değerlendirilmiştir.

Gebelik haftaları ile gebelik sonuçları (normal doğum, sezaryen, terminasyon) arasındaki ilişki Pearson Ki-kare testi ile analiz edilmiştir. Ayrıca, tarama testi varlığı ile QF-PCR sonuçları arasındaki ilişki de Pearson Ki-kare testi ile değerlendirilmiş olup, p-değerleri raporlanmıştır (Tablo 7.7).

Çalışmada, QF-PCR, karyotipleme ve array test sonuçları arasındaki tanısal uyum, Ki-kare testi ve Spearman korelasyon katsayıları kullanılarak incelenmiştir. Bu analizler, prenatal tanıda farklı genetik testlerin sonuçlarının karşılaştırmalı değerlendirilmesini sağlamaktadır.

Tüm istatistiksel analizler, R programlama dili kullanılarak gerçekleştirilmiştir. Veri analizi ve grafiksel gösterimler için temel R paketleri kullanılmıştır.

## 5.6 Dağılımlara Yönelik Tamamlayıcı Analizler

### 5.6.1 Yaş Verisinin Görselleştirilmesi ve QF-PCR Sonuçlarıyla İlişkisi

Yaş dağılımının görselleştirilmesi için yoğunluk grafiği oluşturulmuştur. Veriler, R programlama dilinde `ggplot2` paketi kullanılarak işlenmiş ve yaşların yoğunluğu `geom_density` fonksiyonu ile analiz edilmiştir. Şekil 5.2'te, yaş verilerinin genel dağılımını göstermek için kullanılan R kodu verilmiştir. Yaş verilerinin dağılımını etkili bir şekilde analizi ve görselleştirmesi Şekil 7.1'de verilmiştir.

```r
library(ggplot2)
yas_data <- read.csv("yas_verileri.csv")

# Veri çerçevesi oluşturma
data <- data.frame(Yas = yas)

# ggplot ile yoğunluk grafiği
ggplot(data, aes(x = Yas)) +
  geom_density(fill = "blue", alpha = 0.5) +
  labs(title = "Yaş Dağılımı Yoğunluk Grafiği", x = "Yaş", y = "Yoğunluk") +
  theme_minimal()
```

**Şekil 5.2** Yaş dağılımını görselleştirmek için kullanılan R kodu

Aynı işlevler ve yöntemler, yaş verilerinin QF-PCR sonuçlarına göre gruplandırılarak analizi ve görselleştirilmesi için de uygulanmıştır. Yaş dağılımı, QF-PCR sonuçlarına göre gruplanmış ve farklı renkler ile belirtilmiştir. Bu analiz, yaş gruplarındaki normal ve anormal sonuçların görsel karşılaştırmasını sağlamıştır. Şekil 5.3'de sunulan R kodu ile yapılmıştır. Ayrıntılı analiz için `theme_minimal` teması kullanılarak sade ve okunabilir bir biçimde tasarlanmıştır. Genel yaş dağılımının yanı sıra QF-PCR sonuçlarına göre yaş gruplarının yoğunluk analizi Şekil 7.2'de

görselleştirme yoluyla anlamlı karşılaştırmalar yapılmasına olanak tanımıştır.

```r
library(ggplot2)
veri <- read.table(text = veri_metin, header = TRUE, sep = ",")
veri$yas <- as.numeric(veri$yas)
veri$qfpcr_sonuc <- as.factor(veri$qfpcr_sonuc)
print(ggplot(veri, aes(x = yas, fill = qfpcr_sonuc)) +
    geom_density(alpha = 0.7) +
    labs(title = "Yas - QF-PCR Sonucuna Gore Yogunluk Grafigi",
      x = "Yas",
      y = "Yogunluk",
      fill = "QF-PCR Sonuc"
    ) +
    theme_minimal()
)
```

**Şekil 5.3** Yaş ve QF-PCR sonuçlarına göre dağılımı görselleştirmek için kullanılan R kodu

### 5.6.2 USG Bulguları ve QF-PCR Sonuçlarının Tanımlayıcı ve Görselleştirilmiş Analizi

Bu çalışmada ultrasonografi (USG) bulgularının prenatal tanıda QF-PCR sonuçları ile ilişkisi değerlendirilmiştir. USG sonuçları normal ve anormal olarak sınıflandırılmış ve bu sınıflandırma doğrultusunda QF-PCR sonuçları ile ilişkisi incelenmiştir (Tablo 7.8).

Analiz kapsamında, farklı USG kategorilerinde QF-PCR sonuçlarının dağılımı tanımlayıcı istatistikler kullanılarak değerlendirilmiş ve grafiksel analiz yöntemleri ile görselleştirilmiştir (Şekil 7.3). USG bulgularının prenatal tanıda QF-PCR test sonuçları ile olan ilişkisini belirlemek amacıyla, QF-PCR sonuçlarının normal (0) veya anormal (1) olarak dağılımı incelenmiştir.

Bu analiz, farklı USG kategorilerindeki gebelerin QF-PCR test sonuçları açısından nasıl bir dağılım sergilediğini anlamaya yönelik yapılmıştır. Sonuçların yorumlanması sırasında, farklı USG bulgularına sahip vakalar arasındaki QF-PCR test sonuçlarının dağılımı göz önünde bulundurulmuştur.

### 5.6.3 Prenatal Tarama Testlerinin Terminasyon Kararlarına Etkisinin Dağılım Analizi

Bu çalışmada, prenatal tarama testi yaptıran ve yaptırmayan hastaların gebelik sonuçları karşılaştırmalı olarak değerlendirilmiştir. Excel hesaplamaları kullanılarak,

grupların dağılımları hesaplanmıştır.

Veri seti, toplam 354 gebeye ait prenatal tanı başvurularını içermektedir. Analiz kapsamında, tarama testi yaptıran ve yaptırmayan iki grup arasındaki gebelik sonucu dağılımı karşılaştırılmıştır. Tarama testi yaptıran ve yaptırmayanların, gebelik sonucu aşağıdaki kategorilere ayrılmıştır:

- Gebelik devam etti (normal doğum veya sezaryen ile sonlandı),
- Abort (düşük),
- Terminasyon (gebeliğin isteğe bağlı veya fetal anomali nedeniyle sonlandırılması).

Tarama testi yaptıran ve yaptırmayan grupların gebelik sonucu dağılımları Şekil 7.9'de görselleştirilmiştir. Elde edilen sonuçların makine öğrenmesi modeli ile değerlendirmesi ve karşılaştırılması yapılmıştır.

## 5.7 Makine Öğrenmesi Entegrasyonu

Bu çalışmada, prenatal trizomi riskinin tahmini amacıyla dört farklı makine öğrenmesi algoritması incelenmiş ve karşılaştırmalı performans analizi gerçekleştirilmiştir. Çalışmamızda prenatal trizomi riskini tahmin etmek amacıyla Lojistik Regresyon, Random Forest, Karar Ağacı ve KNN makine öğrenmesi algoritmaları kullanılmıştır. Modellerde kullanılan bağımsız değişkenler (demografik, klinik ve genetik veriler), bağımlı değişken olan 'test sonucu' ile ilişkilendirilerek analiz edilmiştir. Modellerin performansı, kullanılan düzenleme teknikleri ile optimize edilmiştir.

Makine öğrenmesi modellerinin güvenilir ve tekrarlanabilir sonuçlar sunabilmesi için veri seti, eğitim (%70) ve test (%30) setlerine bölünmüştür. Bu işlem, veri dengesini koruyarak model performansını nesnel bir şekilde değerlendirmek amacıyla gerçekleştirilmiştir. Şekil 5.4'de yer alan R kodu, veri setinin eğitim ve test setlerine ayrılmasını sağlamış, bu süreçte kullanılan kodun tekrar edilebilirliği ve doğruluğu güvence altına alınmıştır.

İlk olarak, `set.seed(123)` fonksiyonu ile rastgele sayı üreteci sabitlenerek analizin tekrarlanabilirliği sağlanmıştır. Ardından, `createDataPartition()` fonksiyonu kullanılarak veri seti, hedef değişken olan 'QF-PCR Sonucu' temel alınarak %70 eğitim ve %30 test olmak üzere iki parçaya ayrılmıştır. Bu bölümleme işleminde, `list=FALSE` parametresi ile sonuçların düz bir vektör formatında

```r
# Tekrarlanabilirlik için seed değerini ayarlandı.
set.seed(123)

# Veri setini eğitim (%70) ve test (%30) olarak bölündü.
trainIndex <- createDataPartition(df$qfpcr_sonuc, p = 0.7,
                                  list = FALSE,
                                  times = 1)

# Eğitim ve test setlerini oluşturmak için ayrım yapıldı.
trainData <- df[trainIndex, ]  # %70 eğitim verisi
testData <- df[-trainIndex, ]  # %30 test verisi

# Hedef değişkenin dağılımını kontrol etmek adımı
print(table(trainData$qfpcr_sonuc))
print(table(testData$qfpcr_sonuc))
```

**Şekil 5.4** QF-PCR veri setinin makine öğrenmesi için %70 eğitim ve %30 test verisi olarak bölünmesi

elde edilmesi, `times=1` parametresi ile tek bir bölümleme yapılması sağlanmıştır. Oluşturulan 'trainData' ve 'testData' setlerinin dengeli bir şekilde bölünüp bölünmediğini kontrol etmek amacıyla, hedef değişkenin her iki setteki dağılımı `table()` fonksiyonu kullanılarak Şekil 7.4'de gösterildiği gibi doğrulanmıştır. Bu sistematik yaklaşım, algortimanın eğitimi ve değerlendirmesi için güvenilir bir temel oluşturmaktadır. Ayrıca 10 katlı çapraz doğrulama (cross-validation) yöntemi kullanılmış, her döngüde ROC eğrisi çizilmiş ve AUC değerleri hesaplanmıştır.

Random Forest modelinin hiperparametreleri grid search yöntemiyle ve `tune.grid()` fonksiyonu kullanılarak optimize edilmiştir. Bu süreçte ağaç sayısı (`n_estimators`) 100 olarak belirlenmiş ve her ağacın maksimum düğüm derinliği (`max_depth`) 5 olarak, çapraz doğrulama yöntemleri kullanılarak modelin aşırı öğrenme (overfitting) riskini minimize edecek şekilde ayarlanmıştır. Bu sistematik optimizasyon süreci, modelin aşırı öğrenme riskini minimize ederken tahmin performansını maksimize etmeyi amaçlamıştır.

Bu optimizasyon süreci, modelin hem yeterli öğrenme kapasitesine sahip olmasını hem de aşırı öğrenmeyi engelleyecek şekilde dengelenmesini sağlamıştır. Modelin özellik önem derecesi (feature importance) analizi, prenatal trizomi riskinin tahmininde etkili olan değişkenleri belirlemek için gerçekleştirilmiştir. Bu analiz, karar verme süreçlerinde önemli değişkenlerin belirlenmesini kolaylaştırmıştır.

Çalışmamızdaki algoritmaları tercih etme sebeplerimiz:

- **Lojistik Regresyon:** Trizomi tahmini gibi ikili sınıflandırma problemleri için uygun bir temel algoritma olduğu için,
- **Random Forest:** Farklı demografik, klinik ve genetik değişkenlerin katkılarını anlamak için,
- **Karar Ağacı:** Özellikle kolay yorumlanabilirlik ve hızlı sonuç sunması,
- **K-Nearest Neighbors (KNN):** Daha düşük veri setlerinde kullanarak performans karşılaştırması yapmak için kullanılmıştır.

Modellerin performansını değerlendirmek için şu metrikler kullanılmıştır:

- **Doğruluk (Accuracy):** Modellerin doğru tahmin yapma oranını gösterir.
- **F1 Skoru:** Dengeyi sağlamak için kullanılan bir metrik olup, hatalı pozitif ve hatalı negatif tahminler göz önünde bulundurularak hesaplanmıştır.
- **Eğri Altındaki Alan (AUC - Area Under the Curve):** ROC eğrisinin altındaki alanı ölçerek modelin genel sınıflandırma performansını değerlendirilmiştir.
- **Kesinlik (Precision) ve Duyarlılık (Recall):** Özellikle hatalı pozitif ve hatalı negatif sonuçların etkisini ölçmek için bu iki metrik dikkate alınmıştır.

```
# ROC eğrilerinin çizimi ve karşılaştırılması yapıldı.
plot(log_roc, col = "blue", lwd = 2, main = "ROC Curves for Models")
lines(rf_roc, col = "darkgreen", lwd = 2)
lines(tree_roc, col = "red", lwd = 2)

# Random Forest özellik önem derecesinın çıktısını alındı.
importance_rf <- importance(rf_model)
varImpPlot(rf_model)
```

**Şekil 5.5** QF-PCR verilerinin sınıflandırılmasında kullanılan farklı modellerin ROC eğrileri ve Random Forest modeli için özellik önem analizi kodlaması

Model performansı, ROC (Receiver Operating Characteristic-Alıcı İşlem Karakteristiği) eğrilerinin görselleştirilmesi ve özellik önem derecelerinin analizi ile değerlendirilmiştir. Şekil 5.5'de, lojistik regresyon, random forest ve karar ağacı modellerine ait ROC eğrileri karşılaştırmalı olarak sunulmuş ve her modelin farklı eşik değerlerindeki duyarlılık ve özgüllük performanslarını görsel olarak karşılaştırma imkanı sunmaktadır. Model performans değerlendirmesi kapsamında

gerçekleştirilen analizde, ROC eğrilerinin karşılaştırmalı analizi ve özellik önem derecelerinin görselleştirilmesi olmak üzere iki temel yöntem kullanılmıştır. Random Forest algoritmasıyla yapılan özellik önem derecesi analizi, trizomi tahmininde kritik değişkenlerin belirlenmesine olanak sağlamış ve bu bulgular, prenatal tanıda klinik karar süreçlerini destekleyecek bir temel oluşturmuştur.

Makine öğrenmesi modellerinin performans değerlendirmeleri ve görselleştirmeleri R programlama dilinde gerçekleştirilmiştir. ROC eğrilerini görselleştirme işlemi `plot()` ve `lines()` fonksiyonları kullanılarak yapılmıştır. Random Forest modelinin önem derecelerinin (feature importance) görselleştirilmesi `importance()` ve `varImpPlot()` fonksiyonları ile yapılmıştır. Şekil 7.5, 7.6 ve 7.7'te gösterilmiştir.

# 6
# TARTIŞMA

## 6.1 Çalışmanın Klinik Önemi ve Literatür Karşılaştırması

Prenatal tanı süreçleri, fetüslerde dönemde Trizomi (13, 18, 21) gibi kromozomal anomalilerin erken teşhisi, gebelik yönetimi, fetal sağlık sonuçlarının iyileştirilmesi ve aile bilgilendirilmesi açısından kritik önem taşır. Trizomi 21 (Down sendromu), Trizomi 18 (Edwards sendromu) ve Trizomi 13 (Patau sendromu) gibi kromozomal bozuklukların erken ve doğru teşhisi, yalnızca tıbbi sonuçları değil, aynı zamanda ailelerin karar verme süreçlerini de derinden etkilemektedir. Prenatal dönemde hastaya konulan tanıdan sonra yapılan işlemler (QF-PCR testi, ultrason sonucu, tarama testleri vs.) neticesinde bir teşhise karar verilmesi gerekmektedir. Bu süreçlerde geleneksel yöntemlerin sınırları, yapay zeka ve makine öğrenmesi algoritmalarının kullanılmasını giderek daha cazip hale getirmektedir. Özellikle NIPT ile elde edilen hücre dışı fetal DNA analizine dayalı veri setlerinin genişlemesi, bu teknolojilerin potansiyelini daha da artırmaktadır [80, 81].

Makine öğrenmesi ile süreci optimize etme noktasında dikkat edilmesi ve uygulanması gereken adımlar mevcuttur. Bunların başında detaylı anamnez alınması; yaş gebelik sayısı, aile öyküsü, önceki gebelik sonuçları gibi temel demografik ve klinik verilerin standart bir formatta toplanması gerekmektedir. Ultrason tarama bulguları, fetal ense kalınlığı, burun kemiği ölçümlerinin doğru ve standardize bir şekilde kaydedilmesi gerekir. Biyobelirteçlerden serbest beta-hCG, PAPP-A ve inhibin A gibi analizlerin de standardizasyonu sağlanmalıdır.

Jamshidnezhad ve arkadaşlarının çalışmasında %99.8 doğruluk elde edilirken, bizim çalışmamızda Random Forest modelinin %99 AUC değerine ulaşması, yöntemin tutarlılığını desteklemektedir [82]. Bununla birlikte, çalışmamızda ek olarak USG verilerinin entegrasyonu ile klinik uygulanabilirlik artırıldı.

Geleneksel serolojik testler ve ultrason bulgularına dayanan prenatal risk değerlendirme yöntemleri, düşük maliyet ve geniş erişilebilirlik avantajları

sunmasına rağmen sınırlı doğruluğa sahiptir. Örneğin, bu yöntemlerde %5 yanlış pozitif oranı ve %60-65 arasında değişen tespit oranları rapor edilmektedir. Bu tür sonuçlar, gereksiz invaziv testlerin yapılmasına ve buna bağlı komplikasyon riskinin artmasına yol açabilmektedir [82, 83]. Buna karşılık, NIPT gibi yöntemler, %99'a ulaşan doğruluk oranları ile invaziv prosedürlere alternatif olarak görülmektedir. Ancak, bu testlerin yüksek maliyetleri ve mikrodeletion gibi daha küçük genetik anomalileri tespit etmekteki sınırlamaları, daha verimli ve erişilebilir çözümler arayışını sürdürmektedir [81, 84].

Yapay zeka ve makine öğrenmesi algoritmaları, mevcut yöntemlerin sınırlamalarını aşmak için yenilikçi bir çözüm sunmaktadır. Özellikle Random Forest ve Destek Vektör Makineleri (SVM) gibi algoritmalar, geniş veri kümelerinden öğrenme yetenekleri sayesinde tahmin doğruluğunu artırmaktadır. Çalışmalar, RF algoritmasının Trizomi 21 tahmininde AUC değerini 0.99'a kadar yükselttiğini ve F1 skoru ile yüksek doğruluk oranları sağladığını göstermektedir [80]. Bununla birlikte, SVM algoritmaları, özellikle veri sınıflarının dengesiz olduğu durumlarda daha iyi performans gösterebilmektedir. Bu algoritmaların klinik uygulamalara entegrasyonu, yalnızca doğru teşhis sağlamayı değil, aynı zamanda gereksiz testlerin önlenmesini de kolaylaştırabilir [83].

Genetik verilerin prenatal tanı süreçlerindeki önemi giderek artmaktadır. Jamshidnezhad ve arkadaşları, genetik algoritmalar kullanarak Trizomi 21 risk tahmininde %99.8 doğruluk elde etmişlerdir [82]. Literatürde, genetik ve klinik verilerin kombine edilmesinin tahmin doğruluğunu artırdığı belirtilmişti [81]. Çalışmamızda kullanılan Random Forest algoritması, hem genetik hem de demografik verileri işleyerek çok yönlü bir tahmin modeli sunmaktadır.

Derin öğrenme ve sinir ağı tabanlı modeller, geleneksel yöntemlerin ötesine geçerek daha karmaşık ilişkileri yakalama kapasitesine sahiptir. Çok katmanlı sinir ağları (DNN), mevcut ticari yazılımlardan daha iyi sonuçlar sunarak, %78 tespit oranları ile yanlış pozitif oranlarını azaltmaktadır [83]. Bunun yanı sıra, gen ekspresyon profiline dayalı biyobelirteç panelleri, Trizomi 21 gibi anomalilerin teşhisinde umut vaat eden sonuçlar elde etmektedir. Örneğin, biyobelirteç bazlı modellerin AUC değerleri 1.00'e kadar çıkmakta ve bu durum, genetik verilerin klinik uygulamalarda kullanılabilirliğini artırmaktadır [85]. Çalışmamızda kullanılan veri seti (354 hasta) ve model kombinasyonları, literatürdeki bu bulgularla uyumludur.

## 6.2 Metodolojik Değerlendirme ve Model Performansı

Çalışmamızda QF-PCR testleri, genetik bozuklukları hızlı ve güvenilir bir şekilde tespit edebilme kapasitesi sunarken, makine öğrenmesi algoritmalarının (Lojistik Regresyon, Random Forest, Karar Ağacı, K-Nearest Neighbors) entegrasyonu ile elde edilen sonuçlar trizomi riskini öngörmede büyük bir potansiyel sunmaktadır. Özellikle Random Forest algoritması, 0.99 AUC değeri ile trizomi tahmininde en başarılı model olarak öne çıktı.

Literatür ile karşılaştırıldığında, modelimizin özgün katkısı daha net ortaya çıkmaktadır. Jamshidnezhad ve arkadaşlarının çalışması %99.8 doğruluk oranı ile yüksek bir başarı göstermekle birlikte, yalnızca genetik verilere odaklanmıştır. Bizim çalışmamız ise genetik verilerin yanı sıra demografik ve klinik bilgileri de entegre ederek daha kapsamlı bir yaklaşım sunmaktadır. Benzer şekilde, Sládeček ve arkadaşlarının QF-PCR sonuçlarını analiz ettiği çalışmada sadece QF-PCR verileri kullanılarak %95 doğruluk oranına ulaşılmıştır. Bu karşılaştırmalar, klinik ve demografik verilerin entegrasyonunun model performansına olan katkısını açıkça ortaya koymaktadır. Çalışmamız, makine öğrenmesi tabanlı modellerin klinik karar süreçlerine etkisini geniş bir perspektiften ele alan öncü çalışmalardan biri olma özelliği taşımaktadır.

Çalışmamızda geliştirilen modelin klinik uygulamalara entegrasyonu için şu adımlar önerilebilir:

1. Sağlık Personeline Yönelik Eğitim Programları: Model çıktılarının doğru yorumlanabilmesi ve klinik süreçlere uygun şekilde entegre edilebilmesi için sağlık çalışanlarına yönelik eğitimlerin planlanması gereklidir.
2. Modelin Klinik Karar Destek Sistemlerine Entegrasyonu: Makine öğrenmesi algoritmalarının, hastane bilgi sistemleri ile uyumlu hale getirilmesi, klinisyenlerin karar verme süreçlerini hızlandırabilir.
3. Gerçek Zamanlı Veri Akışı İçin Teknik Altyapı: Hastane sistemlerinde gerçek zamanlı veri akışını sağlayacak altyapıların oluşturulması, modelin etkin bir şekilde kullanılmasını mümkün kılacaktır.

Bu tür karşılaştırmalı analizlerin klinik karar süreçlerine entegrasyonu, tanı ve tedavi yaklaşımlarının etkinliğini artırabilir.

Makine öğrenmesi tabanlı yaklaşımlar, yalnızca doğruluğu artırmakla kalmayıp aynı zamanda ekonomik ve etik açılardan da faydalar sunmaktadır. DNA tabanlı tarama

testleri, cffDNA testlerine kıyasla daha düşük maliyetli bir alternatif olarak öne çıkmaktadır. Özellikle gelişmekte olan ülkelerde, düşük maliyetli testlerin geniş çapta benimsenmesi, prenatal tanı süreçlerinin daha kapsayıcı hale gelmesine katkı sağlayabilir [84]. Ancak, bu tür genetik bilgilerin kullanımına dair etik kaygılar ve veri güvenliği konuları dikkatle ele alınmalıdır. Ailelerin bu süreçte bilinçli kararlar alabilmesi için genetik danışmanlık hizmetlerinin güçlendirilmesi önem taşımaktadır [81].

Bu çalışmada kullanılan veri setleri, hassas kişisel ve klinik bilgiler içerdiğinden, veri gizliliği ve hasta hakları konusundaki etik sorumluluklar titizlikle gözetildi. Makine öğrenmesi tabanlı tahmin modellerinin klinik uygulamalara entegrasyonu sürecinde ortaya çıkan en önemli etik soru, algoritmaların karar verme süreçlerindeki rolü ve sorumluluğun paylaşımıdır. Özellikle algoritmanın önerdiği sonuçlara dayanan klinik kararların nihai sorumluluğunun nasıl belirleneceği, sağlık hukuku ve etik açısından detaylı değerlendirme gerektiren bir konudur.

## 6.3 Kısıtlılıklar ve Gelecek Perspektifi

Gelecek perspektifi üç temel boyutta değerlendirilebilir: İlk olarak, daha geniş ve çeşitli veri setlerinin kullanımı, model performansını artırmanın yanı sıra, farklı coğrafi bölgelerdeki hasta popülasyonlarının daha iyi temsil edilmesini sağlayacaktır. Bu yaklaşım, modelin genellenebilirliğini ve güvenilirliğini artıracaktır. İkinci olarak, derin öğrenme gibi gelişmiş algoritmaların entegrasyonu, verilerdeki karmaşık örüntülerin tespit edilmesini mümkün kılabilir. Ancak bu tür ileri düzey yöntemlerin, daha yüksek hesaplama gücü ve daha büyük veri setleri gerektirdiği göz önünde bulundurulmalıdır. Son olarak, geliştirdiğimiz metodolojinin klinik karar destek sistemlerine entegrasyonu, prenatal tanının ötesinde, diğer genetik hastalıkların erken teşhisinde de önemli bir paradigma değişikliği yaratma potansiyeli taşımaktadır. Bu entegrasyon, sağlık hizmetlerinin kalitesini ve erişilebilirliğini artırırken, karar verme süreçlerini de daha sistematik hale getirebilir.

Çalışmamızın bazı kısıtlılıkları bulunmaktadır:

- Veri setinin tek merkez deneyimine dayanması
- Nadir görülen kromozomal anomalilerin sınırlı temsili
- Prospektif validasyon ihtiyacı
- Biyokimsayal tarama verilerinin detaylandırılarak eklenmesi ile ML algoritmamız daha iyi bir seviyeye çıkarılabilir ve kısıtlılıklar ortadan

kaldırılabilir.

Gelecek çalışmalarda:

- Çok merkezli veri setleriyle modelin validasyonu
- Farklı etnik grupları içeren daha geniş popülasyonlarda test edilmesi
- Real-time klinik uygulamalarda performans değerlendirmesi hedeflenmelidir.

Geliştirdiğimiz makine öğrenme algoritması, invaziv testlere olan ihtiyacı azaltarak önemli bir maliyet etkinliği sağlama potansiyeline sahiptir. Mevcut durumda 354 hastadan 312'sine uygulanan invaziv testlerin maliyeti hasta başı ortalama 2,362 TL (toplam 736,893 TL) olup, bunların %67'sinin (208 test) gereksiz olduğu tespit edilmiştir. Modelimizin üstün performans metrikleri (AUC: 0.99, F1 skoru: 0.87, Doğruluk:0.80, Kesinlik: 0.83, Duyarlılık: 0.91) ve klinik karar verme süreçlerindeki çeşitli faktörler göz önüne alındığında, yıllık maliyetin 294,757 TL ile 331,602 TL arasında (%40-45) azaltılabileceği öngörülmektedir. Bu tahmin, modelimizin yüksek doğruluk oranlarına rağmen, klinik pratikte karşılaşılabilecek değişkenlikler ve hasta-hekim tercihlerini de dikkate alan ihtiyatlı bir yaklaşımı yansıtmaktadır.

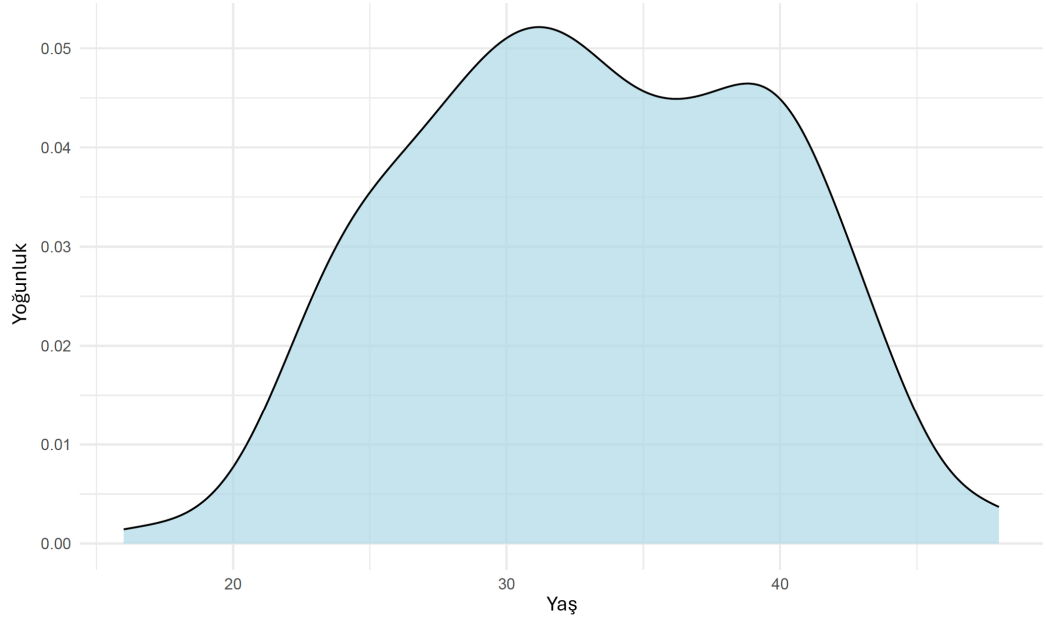
Yapay zeka ve makine öğrenmesi destekli modeller, prenatal tanı süreçlerinde devrim yaratma potansiyeline sahiptir. Bu teknolojiler, yalnızca doğruluk oranlarını artırmakla kalmayıp, invaziv prosedürlere olan ihtiyacı da azaltmaktadır. Daha fazla klinik çalışma ve model optimizasyonu ile bu yaklaşımların yaygınlaştırılması, gebelik yönetimini önemli ölçüde iyileştirebilir. Ayrıca, bu modellerin halk sağlığı politikalarına entegrasyonu, geniş popülasyonların sağlık sonuçlarını optimize etme potansiyeline sahiptir. Gelecekte, genetik ve klinik verilerin entegrasyonunu daha ileri taşıyacak yeni nesil yapay zeka araçlarının geliştirilmesi, prenatal tanı süreçlerinin dönüştürülmesinde kilit rol oynayacaktır.

# 7
# SONUÇ

## 7.1 Demografik, Klinik ve Genetik Verilerin Analizi

### 7.1.1 Popülasyon Analizi ve Yaş Dağılımı

Çalışmada kullanılan demografik veriler, genetik risk faktörlerini belirlemek için önemli rol oynamıştır. Çalışmaya dahil edilen 354 hastanın demografik ve klinik özellikleri detaylı olarak analiz edilmiştir. Yaş dağılımı incelendiğinde, hastaların %48.3'ünün (n=171) 31-40 yaş aralığında olduğu, bunu %35.6 (n=126) ile 21-30 yaş grubu ve %14.4 (n=51) ile 41-50 yaş grubunun izlediği, %1.42'sinin (n=5) ise 0-20 yaş aralığında olduğu görülmüştür. Çalışma popülasyonunun yaş ortalaması $33.06 \pm 6.62$ yıl olarak hesaplanmıştır. Erkek olduğu tespit edilen 1 hasta çalışmaya dahil edilmemiştir.




**Şekil 7.1** Popülasyonun yaş dağılımını gösteren yoğunluk grafiği

Yaş dağılımı yoğunluk grafiği Şekil 7.1'de, çalışma popülasyonundaki yaşların büyük ölçüde 20 ile 40 yaş aralığında yoğunlaştığını ve 40 yaş üzerinde yoğunluğun azaldığını ortaya koymaktadır. Yoğunluk eğrisi, popülasyonun demografik yapısını

görsel olarak ifade etmekte ve önemli bir genel bakış sunmaktadır. Yoğunluk grafiğinde tepe noktalarının belirgin olduğu görülmektedir. Bu durum, popülasyonda yaşların belirli aralıklarda yoğunlaştığını ve bu aralıkların modal değerleri temsil ettiğini göstermektedir. Örneğin, 20-30 yaş ve 30-40 yaş aralığında belirgin bir yoğunluk gözlenirken, 40 yaşın üzerinde yoğunluğun azaldığı görülmektedir. Bu durum, popülasyonun büyük bir kısmının genç ve orta yaş gruplarından oluştuğunu, yaşlı bireylerin ise popülasyonda daha az temsil edildiğini işaret etmektedir.

### 7.1.2 Klinik Bulgular ve Gebelik Özellikleri

Gebelik özellikleri değerlendirildiğinde, hastaların çoğunluğunun (%75.1, n=266) çoklu gebelik öyküsü olduğu, %24.0'ünde (n=85) ise ilk gebeliği olduğu tespit edilmiştir. Popülasyonda ortalama gebelik sayısı $2.75 \pm 1.69$, ortalama doğum sayısı ise $1.21 \pm 1.18$ olarak bulunmuştur. Klinik özellikler incelendiğinde, hastaların %56.2'sinde (n=199) ultrasonografi anomalisi saptanmış, %43.79'unda (n=155) ultrasonografi bulguları normal bulunmuştur. Prenatal tanı endikasyonları arasında ileri anne yaşı (%39.5), anormal USG bulguları (%35.7), tarama testlerinde yüksek risk (%18.3) ve önceki gebelikte kromozom anomalisi öyküsü (%6.5) en sık nedenlerdir.

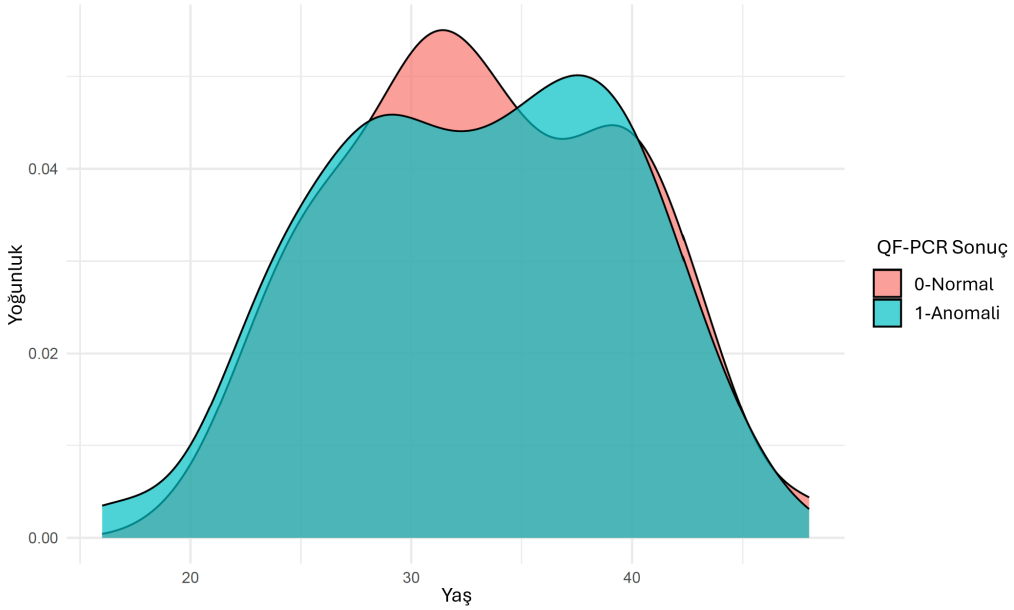
### 7.1.3 Genetik Analiz Sonuçları ve Yaş İlişkisi

Çalışmada kullanılan demografik veriler, genetik risk faktörlerini belirlemek için yardımcı olmuştur. Genetik veriler, tanı doğruluğunu artıran ve trizomi varlığını kesinleştiren verilerdir. Klinik veriler, tarama ve ultrason sonuçları gibi bulgularla risk seviyesi değerlendirilmiştir. QF-PCR sonuçlarına göre, hastaların yaş ortalamaları ve standart sapmaları elde edilmiştir. Sonuçlar Tablo 7.1'de gösterildiği gibi detaylandırılmıştır. Elde edilen QF-PCR sonuçları, belirli genetik anomalilerin ve trizomilerin yaş dağılımı ile ilişkisini ortaya koymaktadır. QF-PCR analizleri, Trizomi 21'in (%14.4, n=51) en sık gözlenen anomali olduğu ve yaş ortalamasının $34{,}4 \pm 7{,}6$ olduğu görülmüştür. Bunun dışında Trizomi 18 (n=17), Trizomi 13 (n=11), Triploidi (n=3) ve Monozomi X (n=4) gibi daha az sıklıkla görülen anomaliler tespit edilmiştir. Analiz sonuçlarına göre genetik anomalilerin büyük bir kısmı (%90 üzerinde) maternal kalıtım kaynaklıdır. Bu durum, ileri anne yaşı gibi maternal faktörlerin prenatal tanı süreçlerindeki kritik önemini vurgulamaktadır. İleri anne yaşının Trizomi 21 prevalansı üzerindeki etkisi, özellikle 35 yaş ve üzerindeki gebelerde daha belirgin hale gelmektedir. Literatürde yapılan çalışmalar bu yaş grubunda anöploidi riskinin %15-20 arttığını göstermektedir [86, 87]. Çalışmamızda bu durum, Trizomi 21 olgularında gözlemlenen yaş ortalamasının (34,4 ± 7,6 yıl)

yüksek olmasıyla uyumlu bulunmuştur. Bu bulgular, yaş faktörünün prenatal tanıda genetik anomalilerin tahmininde önemli bir belirleyici olduğunu desteklemektedir. Bu bulgular, klinik değerlendirme ve prenatal tanı süreçlerindeki yaş faktörünün önemini vurgulamaktadır.

**Tablo 7.1** QF-PCR sonuçlarının sayıları, yaş ortalaması ve standart sapma

| QF-PCR Sonucu | Sayı | Yaş Ortalaması($\pm$SS) |
|---|---|---|
| XXY | 3 | $30.7 \pm 3.8$ |
| Triploidi | 3 | $25.7 \pm 3.2$ |
| Monozomi X | 4 | $27.5 \pm 4.4$ |
| Anomali | 15 | $30.1 \pm 5.1$ |
| Trizomi 13 | 11 | $29.4 \pm 7.2$ |
| Trizomi 18 | 17 | $33.9 \pm 6.5$ |
| Trizomi 21 | 51 | $34.4 \pm 7.6$ |
| Normal | 250 | $33.3 \pm 6.4$ |




**Şekil 7.2** Yaşa göre QF-PCR sonuçlarının (0-Normal, 1-Anomali) yoğunluk dağılımı

QF-PCR sonuçlarının yaş değişkenine göre yoğunluğu Şekil 7.2'de gösterilmiştir. Verilerde, QF-PCR normal sonuçlar (0) ve anomali sonuçları (1) için ayrı ayrı incelenebilecek dört belirgin tepe noktası bulunmaktadır. Normal QF-PCR sonuçları (kırmızı tonlu renk), genç yaş grubunda ($\sim$27-35 yaş aralığında, 0.05 yoğunluğunu aşan) bir tepe noktasına sahiptir ve bu durum bu yaş grubunda düşük trizomi riski ile uyumludur. İkinci bir tepe noktası ise 35 yaş civarında gözlemlenmektedir. Anomali QF-PCR sonuçları (turkuaz renk) ise yaklaşık 37 yaş ve üzerindeki yaş grubunda yoğunlaşmaktadır. Burada, ileri yaşın trizomi riski üzerinde belirgin bir etkisi olduğu görülmektedir.

Genç yaş grubundaki yoğunluk, bu yaş aralığında normal sonuçların daha yaygın olduğunu gösterirken, ileri yaş grubundaki artış anomali sonuçlarıyla ilişkilendirilmektedir. İki dağılımın çakışma bölgeleri, ileri yaşta dahi normal sonuçların sık görülebildiğini ancak anomali sonuçlarının bu yaş grubunda anlamlı ölçüde arttığını ortaya koymaktadır. Bu grafik, yaşın genetik anomaliler üzerindeki etkisini anlamada önemli veriler sunmakta ve QF-PCR'nin klinik uygulamalarda kritik bir araç olduğunu göstermektedir.

### 7.1.4 Prenatal Tanı Endikasyonları ve Numune Dağılımı

Prenatal tanı için başvuru nedenleri analiz edildiğinde, ileri anne yaşının en sık başvuru nedeni olduğu, bunu sırasıyla anormal ultrasonografi bulguları, tarama testlerinde yüksek risk ve önceki gebelikte kromozom anomalisi öyküsü izlediği görülmüştür. QF-PCR testi için alınan numunelerin dağılımında ise amniyon sıvısı örnekleri %90.68 (n=321) ile en büyük grubu oluştururken, bunu %1.69 (n=6) ile koryon villus biyopsisi örnekleri takip etmiştir. Abort materyali örnekleri %3.39 (n=12) oranında görülürken, diğer numune türleri %4.24 (n=15) oranında tespit edilmiştir.

### 7.1.5 QF-PCR Sonuçları ve Gebelik Sonuçları İlişkisi

QF-PCR sonuçları ile gebelik sonuçları arasındaki ilişki Tablo 7.2'de değerlendirildiğinde, normal saptanan olguların büyük çoğunluğunun (102 olgu) sezaryen(CS) ile doğum yaptığı, bunu normal doğum (47 olgu) ve terminasyon (30 olgu) seçeneklerinin izlediği görülmektedir. Anomaliler içinde en sık rastlanan Trizomi 21 olgularının büyük kısmında (33 olgu) gebelik terminasyonu tercih edilmiştir. Trizomi 18, Trizomi 13 ve diğer anomali gruplarında da terminasyon oranlarının yüksek olduğu dikkat çekmektedir. Bu veriler, prenatal tanıda saptanan kromozomal anomalilerin gebelik sonuçları üzerinde belirleyici bir etkiye sahip olduğunu ve ailelerin gebelik yönetimi konusundaki kararlarını önemli ölçüde etkilediğini göstermektedir.

### 7.1.6 Maternal ve Paternal Kalıtım Analizi

Prenatal tanıda QF-PCR analizi sonuçları değerlendirildiğinde Tablo 7.3'de, kromozomal anomalilerin maternal kalıtım oranının paternal kalıtıma göre belirgin şekilde yüksek olduğu gözlenmiştir. Bu durum özellikle Trizomi 21 vakalarında dikkat çekicidir (maternal: 46, paternal: 2). Benzer şekilde, Trizomi 18 olgularında da maternal kalıtımın baskın olduğu görülmüştür. Bu veriler, maternal yaş ve ilişkili risk faktörlerinin prenatal tanıda önemli bir belirleyici olduğunu

**Tablo 7.2** Prenatal QF-PCR test sonuçlarının gebelik sonlanma şekillerine göre dağılımı

| QF-PCR Sonuçları | Abort | NA | Normal Doğum | Terminasyon | CS |
|---|---|---|---|---|---|
| XXY | 0 | 2 | 0 | 0 | 1 |
| Normal | 9 | 62 | 47 | 30 | 102 |
| Trizomi 21 | 0 | 6 | 3 | 33 | 9 |
| Trizomi 18 | 1 | 2 | 1 | 10 | 3 |
| Trizomi 13 | 0 | 0 | 0 | 9 | 2 |
| Monozomi X | 1 | 0 | 0 | 0 | 3 |
| Anomali | 2 | 0 | 0 | 13 | 0 |
| Triploidi | 2 | 0 | 0 | 1 | 0 |

desteklemektedir. Ayrıca, prenatal risk değerlendirmesinde anne yaşının yanı sıra genetik tarama testlerinin birlikte değerlendirilmesinin tanısal açıdan önemini vurgulamaktadır. Bu sonuçlar, özellikle ileri anne yaşı ve diğer risk faktörleri olan gebeliklerde, detaylı genetik danışmanlık ve prenatal tanı testlerinin gerekliliğini ortaya koymaktadır. Maternal faktörlerin kromozomal anomaliler üzerindeki belirgin etkisi, risk değerlendirmesinde anne yaşı ve maternal öykünün dikkatle ele alınması gerektiğini göstermektedir.

**Tablo 7.3** Kromozomal anomalilerin QF-PCR sonuçlarına göre ebeveynsel kalıtım dağılımı

| QF-PCR Sonucu | Maternal Kalıtım | Paternal Kalıtım |
|---|---|---|
| XXY | 1 | 0 |
| Anomali | 7 | 2 |
| Monozomi X | 2 | 0 |
| Triploidi | 3 | 0 |
| Trizomi 13 | 7 | 4 |
| Trizomi 18 | 16 | 1 |
| Trizomi 21 | 46 | 2 |

#### 7.1.6.1 Veri Analizi ve Standardizasyon

Veri ön işleme adımları, modelin veriyi daha doğru bir şekilde işlemesini sağlayarak analiz sonuçlarının güvenilirliğini artırmış ve model performansını optimize etmiştir. Verilerin sınıflandırılmasındaki standardizasyon yaklaşımı, kromozomal anormallikler, QF-PCR sonuçlarının sınıflandırılmasında tutarlılık sağlayarak, genetik araştırmaların ve klinik uygulamaların veri-odaklı perspektifini güçlendirmektedir.

## 7.2 İstatistiksel Analiz Sonuçları

Çalışmaya dahil edilen gebelerin demografik ve klinik özellikleri Tablo 7.4'de özetlenmiştir. Çalışma popülasyonunun yaş ortalaması 33.06 ± 6.62 yıl olup, yaş aralığı 16-48 yıl olarak saptanmıştır. Gebelik haftası ortalaması 19.50 ± 3.85 hafta (aralık: 0-34 hafta) olduğu görülmüştür. Tarama testi varlığının yaş gruplarına göre dağılımı Tablo 7.5'de gösterilmiştir ve değişkeni 0-1 arasında kodlanmıştır. Ortalaması 0.38 ± 0.49 olarak hesaplanmıştır. Bu değer, çalışma popülasyonunun yaklaşık %38'inde tarama testi yapıldığını göstermektedir. Bu üç değişkenin bir arada değerlendirilmesi, gebelik sonuçlarıyla ilişkilerinin analizi açısından güçlü bir temel oluşturmakta ve klinik açıdan anlamlı sonuçlar üretme potansiyeli taşımaktadır.

**Tablo 7.4** Prenatal tarama programına katılan gebelerin demografik ve klinik özelliklerinin tanımlayıcı istatistikleri

| | **Ortalama** | **Medyan** | **Standart Sapma** | **Min** | **Max** |
|---|---|---|---|---|---|
| **Yaş** | 33.06 | 33 | 6.62 | 16 | 48 |
| **Gebelik Haftası** | 19.50 | 20 | 3.85 | 0 | 34 |
| **Tarama Testi** | 0.38 | 0 | 0.49 | 0 | 1 |

**Tablo 7.5** Prenatal tarama testi kullanımının yaş gruplarına göre dağılım analizi

| **Yaş Grubu** | **Tarama Testi Yok(%)** | **Tarama Testi Var(%)** |
|---|---|---|
| 0-20 | 4 | 1 |
| 21-30 | 94 | 32 |
| 31-40 | 101 | 70 |
| 41-50 | 20 | 31 |

Çalışmamızda, prenatal tarama testi varlığı ile gebelik sonuçları arasında istatistiksel olarak anlamlı ilişki saptanmıştır ($p<0.05$). Bu durum yaş grupları ile değerlendirildiğinde (Tablo 7.6'de); ileri maternal yaş gruplarında (41-50 yaş) tarama testi yapılma oranı ve sezaryen doğum oranı daha yüksek bulunmuştur. 13-24 gebelik haftasında hem normal doğum hem de terminasyon oranlarının yüksek olduğu gözlenmiştir. Tarama testi yapılan gebeliklerde terminasyon oranının daha yüksek olması, testlerin trizomi riski tespitindeki etkinliğini desteklemektedir. Bu bulgular, prenatal tanı ve tarama testlerinin özellikle ileri maternal yaşta gebelik yönetiminde önemli rol oynadığını göstermektedir.

Yapılan analizler, prenatal tanıda tarama testlerinin ve QF-PCR sonuçlarının klinik sonuçlarla ilişkisini anlamak için çeşitli istatistiksel yöntemleri kapsamaktadır. Tarama testi varlığı ile gebelik sonuçları arasındaki ilişki analiz anlamlı bulunmamıştır ($p > 0.05$). Ancak, tarama testi varlığı ile QF-PCR sonucu arasındaki ki-kare testi (p-değeri = 0.017) istatistiksel olarak anlamlı bir ilişki

olduğunu göstermiştir. Bu durum, tarama testlerinin varlığının QF-PCR sonuçlarını etkileyebileceğini veya bu sonuçlarla ilişkili olabileceğini göstermektedir.

**Tablo 7.6** Maternal yaş gruplarına göre gebelik sonlanma şekillerinin dağılımı

| Yaş Grubu | CS | Normal Doğum | NA | Terminasyon | Abort |
|---|---|---|---|---|---|
| 0-20 | 0 | 0 | 4 | 1 | 0 |
| 21-30 | 39 | 17 | 22 | 41 | 7 |
| 31-40 | 60 | 24 | 38 | 42 | 7 |
| 41-50 | 21 | 10 | 8 | 12 | 1 |

Yaş gruplarına göre yapılan değerlendirmelerde, sezaryen oranlarının özellikle ileri yaş gruplarında (31-40 ve 41-50) daha yüksek olduğu görülmektedir. Ayrıca, gebelik haftası ve gebelik sonuçları arasındaki ilişki incelendiğinde, terminasyonların daha çok ikinci trimesterde (13-24 hafta) gerçekleştiği gözlenmiştir. Bu da, erken tanının klinik yönetimde kritik bir rol oynadığını ortaya koymaktadır.

QF-PCR sonuçlarını değerlendirmek için karyotipleme ve array testleri arasında yapılan ki-kare analizi, istatistiksel olarak anlamlı bir ilişki bulunmadığını göstermiştir (Tablo 7.7). İlk olarak, QF-PCR ve karyotipleme arasındaki analizde p-değeri 0.726 bulunmuş; bu sonuç, iki test arasında anlamlı bir ilişki olmadığını göstermiştir. Benzer şekilde, QF-PCR ile array arasında yapılan analizde p-değeri 0.986 bulunmuş; bu da yine istatistiksel olarak anlamlı bir ilişki olmadığını göstermiştir.

**Tablo 7.7** Prenatal kromozomal anomali testlerinin (QF-PCR, Karyotipleme ve Array) tanısal uyum ve korelasyon analizi

| Testler | Ki-Kare Değeri | p-Değeri | Spearman Korelasyon Katsayısı | Uyum Oranı (%) |
|---|---|---|---|---|
| QF-PCR ve Karyotipleme | 0.12 | 0.726 | -0.025 | 48.8 |
| QF-PCR ve Array | 0.027 | 0.986 | 0.008 | 33.6 |

Korelasyon analizi sonuçlarına göre, QF-PCR ile karyotipleme arasında Spearman korelasyon katsayısı -0.025 olarak belirlenmiştir. Bu değer, iki test arasında oldukça zayıf ve negatif bir ilişki olduğunu göstermektedir. QF-PCR ve array arasındaki Spearman korelasyon katsayısı ise 0.008 olarak bulunmuş, bu da oldukça zayıf bir pozitif ilişkiyi işaret etmektedir.

Son olarak, uyum oranları incelendiğinde, QF-PCR ile karyotipleme arasında yaklaşık %48.8'lik bir uyum oranı gözlemlenmiştir. Bu, QF-PCR ve karyotipleme sonuçlarının aynı kategoride yer aldığı durumların %50 civarında olduğunu göstermektedir. Öte yandan, QF-PCR ile array arasındaki uyum oranı daha düşük olup yaklaşık %33.6 olarak tespit edilmiştir. Bu sonuçlar, QF-PCR'nin karyotipleme

ve array test sonuçlarıyla güçlü bir paralellik göstermediğini ve sonuçların birbirinden bağımsız olduğunu ortaya koymaktadır.

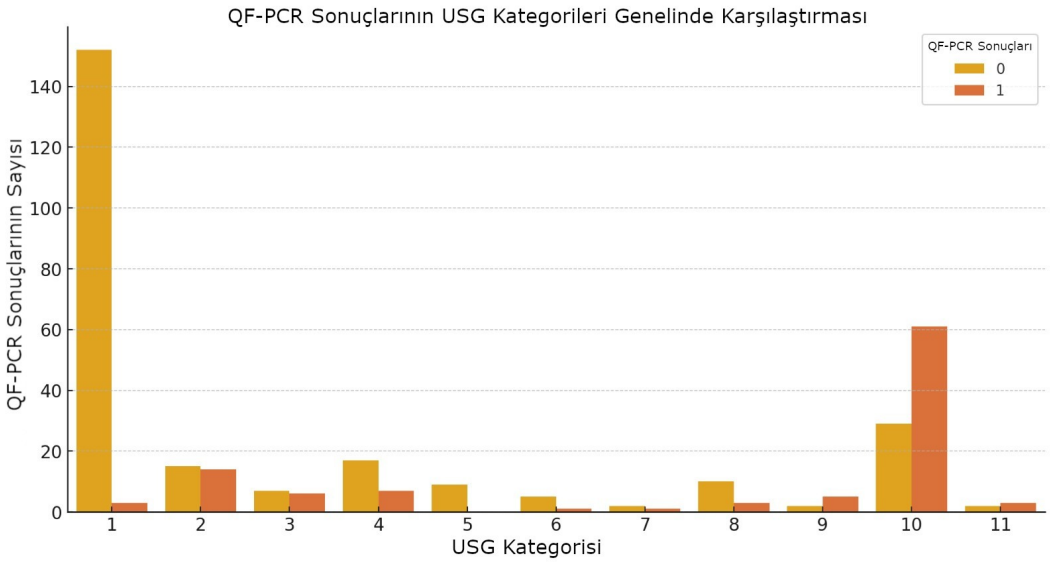
Uyum oranlarını değerlendirirken bazı kısıtlılıkları göz önünde bulundurmak gerekir. Array analizi tüm hastalara uygulanmamış, karyotipleme testi ise çoğunlukla anormal vakaları tespit etmek için kullanılmıştır. Bu nedenle, QF-PCR ile diğer yöntemler arasındaki uyum oranları gerçekte olduğundan daha düşük görünebilmektedir.

Sağlık alanındaki çalışmalar sırasında elde edilen veriler, farklı teşhis ve değerlendirme yöntemlerinin etkinliğini karşılaştırmak açısından büyük önem taşımaktadır. Bu kapsamda, ultrason görüntüleme (USG) kategorileri ile gerçekleştirilen kantitatif floresan polimeraz zincir reaksiyonu (QF-PCR) sonuçlarının incelenmesi önemli bir analiz konusudur. Çalışmanın sonuçlarına göre, ultrasonografi (USG) verilerinin değerlendirmesinde %56.21'i anormal, %43.79'u normal olarak tespit edilmiştir (Tablo 7.8). Yaş gruplarına göre tarama testi varlığı incelendiğinde, 31-40 yaş grubunda test varlığının en yüksek olduğu (n=70) ve test yokluğunun da yine aynı yaş grubunda en fazla olduğu (n=101) gözlenmiştir. Gebelik haftası ve sonuçları değerlendirildiğinde, 13-24 hafta aralığında normal doğum (n=51), sezaryen (n=109), terminasyon (n=88) ve abort (n=15) vakalarının gerçekleştiği belirlenmiştir. Yaş gruplarına göre gebelik sonuçları incelendiğinde, hem terminasyonun (n=42) hem de sezaryen doğumların (n=60) en yüksek oranda 31-40 yaş grubunda görüldüğü saptanmıştır.

**Tablo 7.8** Prenatal ultrasonografi bulgularının normal ve anormal olarak yüzdesel dağılımı

| **Kategori** | **Yüzde (%)** | **Hasta Sayısı** |
|---|---|---|
| Anormal | 56.21 | 200 |
| Normal | 43.79 | 154 |

Şekil 7.3'te grafikte USG bulgularına göre belirlenen 11 farklı kategoride QF-PCR sonuçlarının dağılımı gösterilmektedir. Sarı sütunlar normal QF-PCR sonuçlarını (0), turuncu sütunlar ise anomali olan QF-PCR sonuçlarını (1) temsil etmektedir. Grafikte Y ekseni QF-PCR sonuçlarının vaka sayısını, X ekseni USG kategorilerini göstermektedir. Grafiğe göre, Kategori 1'de normal QF-PCR sonuçları belirgin şekilde yüksek bulunmuştur. Orta kategorilerde (2-3) normal ve anomali sonuçları dengeli bir dağılım gösterirken, yüksek kategorilerde (4-10) anomali oranının kademeli olarak arttığı, özellikle Kategori 10'da en yüksek seviyeye ulaştığı gözlenmiştir. Bu veriler, USG kategorileri ile QF-PCR sonuçları arasında anlamlı bir ilişki olduğunu göstermiştir.




**Şekil 7.3** Ultrason kategorilerine göre normal ve anormal QF-PCR sonuçlarının dağılım analizi

## 7.3 Makine Öğrenmesi Modellerinin Sonuçları

### 7.3.1 Veri Hazırlama ve Normalizasyon

Çalışmada yapılan kategorik sınıflandırma işlemleri, klasik normalizasyon tekniklerinden (örneğin, min-max veya z-score) farklı olmakla birlikte, veri analizi ve modelleme süreçleri için gerekli olan sistematik bir yapı sağlamıştır. Metot bölümünde de ayrıntılı olarak açıklandığı gibi, sürekli değişkenlerde z-skor dönüşümü, belirli bir aralıkta olması gereken değişkenlerde ise min-max normalizasyonu uygulanmıştır. Örneğin, yaş, gebelik haftası gibi sürekli değişkenler z-skor dönüşümü ile ölçeklendirilmiş, anne yaşı gibi belirli bir aralıkta olan değişkenler ise min-max normalizasyonu ile 0-1 aralığında ifade edilmiştir. Ayrıca, büyük sayısal değerlerin daha kolay işlenebilir hale getirilmesi için ondalık ölçekleme kullanılmıştır.

### 7.3.2 Kategorik Değişkenlerin Kodlanması

Kategorik verilerde ise, standartlaştırma ve kodlama işlemleri gerçekleştirilmiştir. Kategorik değişkenler metin formatında standardize edildikten sonra, hem label encoding hem de one-hot encoding teknikleri ile analizlere uygun hale getirilmiştir. Örneğin, gebelik sonuçları ve numune türleri gibi kategorik değişkenler, belirli gruplar altında kodlanarak analiz ve makine öğrenmesi modellerine entegre edilmiştir. Bu süreç, çalışmanın makine öğrenmesi algoritmalarında kullanılan kategorik kodlama altyapısını oluşturmuş ve analizlerin güvenilirliğini artırmıştır.

Ayrıca, aşırı uç değerlerin tespiti ve işlenmesi gibi normalizasyon süreçlerini

destekleyen ek adımlar, istatistiksel analizlerin doğruluğunu sağlamıştır. Örneğin, anne yaşı için belirlenen 16-48 yaş sınırları dışındaki değerler, klinik değerlendirme ile doğrulanarak analiz dışı bırakılmıştır. Tüm bu işlemler, hem istatistiksel analizler hem de makine öğrenmesi modelleri için yapılandırılmış bir veri formatının oluşturulmasını sağlamış, sınıflandırma algoritmaları, desen tanıma ve kümeleme analizleri için uygun bir ön işleme örneği sunmuştur.

Bu kapsamda, çalışmada uygulanan veri düzenleme ve normalizasyon işlemleri, metot bölümünde belirtildiği gibi, veri analizi için daha uygun bir format oluşturmuş, kategorik verilerin kodlanması ve sürekli değişkenlerin ölçeklendirilmesiyle makine öğrenmesi modellerinin etkinliğini artırmıştır.

### 7.3.3 Eğitim ve Test Veri Setleri

Veri seti, eğitim ve test setlerine sırasıyla %70 ve %30 oranında ayrılmıştır. Eğitim ve test setinde hedef değişkenin ("QF-PCR Sonuç") dağılımı Şekil 7.4'de gösterilmiştir. Bu dağılıma göre, eğitim setinde 175 adet "0", normal sınıfı ve 73 adet "1", anomali sınıfı olarak gözlemlenmiştir. Burada "0" sınıfı normal sonuçları, "1" sınıfı ise anomali olan sonuçları ifade etmektedir. Test setinde ise 75 adet "0", normal sınıfı ve 31 adet "1", anomali sınıfı bulunmaktadır. Bu durum, veri setinin eğitim ve test aşamalarında hedef değişkenin sınıf dağılımının dengeli bir şekilde temsil edildiğini göstermektedir. Sınıf oranlarının dengeli bir şekilde ayrılması, makine öğrenimi modellerinin hem normal sonuç hem de anomali sonuç durumları için doğru bir şekilde öğrenme gerçekleştirmesini sağlamaktadır.

```
# Eğitim ve Test Setlerinin Ayrılmış Halleri
> table(trainData$qfpcr_sonuc)
  0   1
175  73
> table(testData$qfpcr_sonuc)
 0  1
75 31
```

**Şekil 7.4** Eğitim ve test veri setlerinin QF-PCR sonuçlarına göre dağılımı (0: Normal, 1: Anomali)

### 7.3.4 Makine Öğrenmesi Modellerinin Karşılaştırmalı Performans Analizi

Model performanslarının karşılaştırmalı analizi, farklı sınıflandırma algoritmalarının çeşitli metrikler açısından değişken başarı düzeyleri sergilediğini Tablo 7.9'da ayrıntılı olarak verilmiştir. Random Forest modelinin, hem AUC (Eğri Altında Kalan Alan) değeri hem de genel doğruluk açısından diğer modellere kıyasla üstün

performans gösterdiği gözlemlenmiştir. Bu sonuç, modelin sınıflandırma görevinde dengeli ve güvenilir bir yaklaşım sunduğunu göstermektedir. Bu çalışma, prenatal tanıda makine öğrenmesi algoritmalarının etkili bir şekilde uygulanabilir olduğunu göstermiştir. Elde edilen bulgular, klinik karar alma süreçlerini güçlendirmek için önemli bir referans sunmaktadır.

**Tablo 7.9** QF-PCR sonuçlarının sınıflandırılmasında %70 eğitim-%30 test veri seti ayrımı ile uygulanan makine öğrenmesi algoritmalarının performans metrikleri

| Method | AUC | F1 | Doğruluk | Duyarlılık | Kesinlik |
|---|---|---|---|---|---|
| Lojistik Regresyon | 0.86 | 0.86 | 0.81 | 0.84 | 0.89 |
| Random Forest | 0.99 | 0.87 | 0.80 | 0.91 | 0.83 |
| Karar Ağacı | 0.83 | 0.88 | 0.82 | 0.92 | 0.84 |
| K-Nearest Neighbors (KNN) | 0.52 | 0.80 | 0.69 | 0.91 | 0.72 |

Random Forest modeli, 0.99 AUC değeri ve 0.87 F1 skoru ile sınıflandırma görevinde üstün performans sergilemiştir. Bu metrikler, modelin hem pozitif-negatif sınıf ayrımında hem de genel sınıflandırma başarısında dengeli ve güçlü sonuçlar elde ettiğini göstermektedir. Bu değer, modelin hem doğru pozitifleri tespit etme hem de yanlış pozitifleri minimize etme konusunda etkili olduğunu kanıtlamaktadır. Genel doğruluk oranının %80 olması, Lojistik Regresyon modeline kıyasla daha yüksek bir başarı elde edildiğini göstermiştir. Duyarlılık metriğindeki %91'luk başarı oranı, modelin pozitif sınıfları tanımlama konusundaki güvenilirliğini ortaya koyarken, %83'lik kesinlik değeri, pozitif sınıf tahminlerindeki yüksek isabetliliği vurgulamaktadır. Bu metrikler, modelin sınıflandırma görevindeki genel başarısını destekler niteliktedir.

Random Forest modelinin en dikkat çekici özelliği, 0.99'luk AUC değeri ile gösterdiği üstün performanstır. Bununla birlikte, modelin birden fazla karar ağacı kullanması nedeniyle yüksek hesaplama maliyeti ve işlem gücü gerektirmesi, pratik uygulamalarda göz önünde bulundurulması gereken bir kısıt olarak değerlendirilmelidir. Bu durum, model seçiminde performans ve kaynak kullanımı arasında bir denge kurulması gerektiğini göstermektedir.

Lojistik Regresyon modelinin performans analizi, çeşitli değerlendirme metrikleri üzerinden kapsamlı bir değerlendirme sunmaktadır. Model, 0.86'lik AUC değeri ile pozitif ve negatif sınıfların ayrıştırılmasında orta düzeyde bir başarı sergilemiştir. Bu sonuç, modelin temel sınıflandırma görevlerinde makul bir performans gösterdiğini işaret etmektedir. Modelin 0.86'lük F1 skoru, kesinlik ve duyarlılık metrikleri arasında dengeli bir performans yakaladığını göstermektedir. Bu değer, modelin doğru pozitifleri tespit etme ve yanlış pozitifleri minimize etme konusunda tutarlı bir yaklaşım sergilediğini ortaya koymaktadır. %81'lik genel doğruluk oranı, modelin

veri setindeki örneklerin önemli bir kısmını doğru şekilde sınıflandırabildiğini göstermektedir. Duyarlılık metriğinde elde edilen %84'lük başarı oranı, modelin pozitif sınıfları tanımlama konusundaki yetkinliğini vurgularken, %89'lik kesinlik değeri, pozitif sınıf tahminlerindeki güvenilirliği desteklemektedir. Bu metrikler bir arada değerlendirildiğinde, modelin dengeli ve tutarlı bir performans sergilediği görülmektedir.

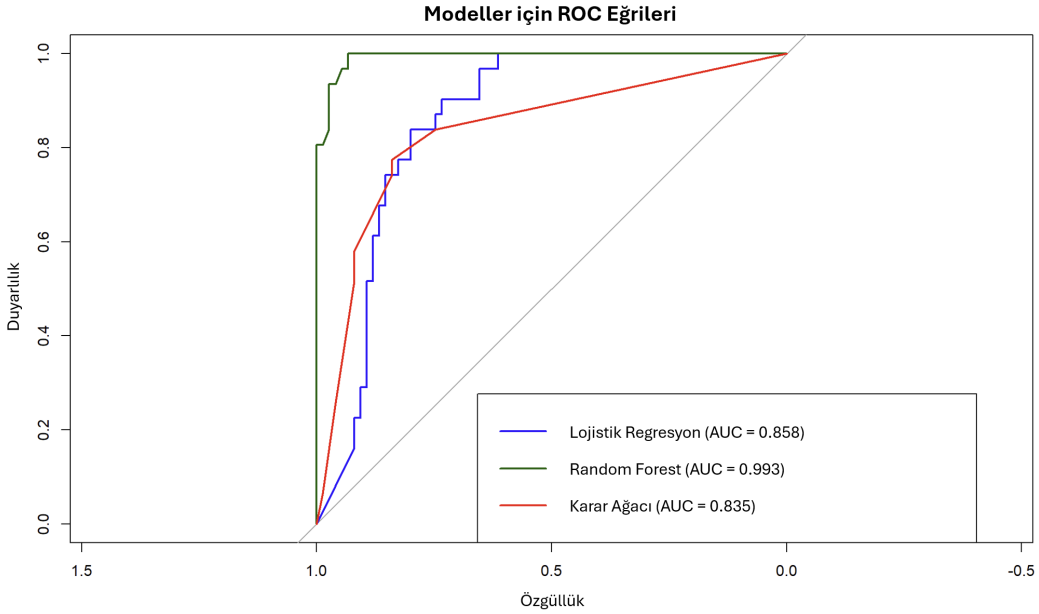
Lojistik Regresyon modelinin en önemli avantajı, düşük hesaplama maliyeti ve hızlı sonuç üretme kabiliyetidir. Bu özelliği, modeli özellikle hızlı prototipleme ve gerçek zamanlı uygulamalar için uygun bir seçenek haline getirmektedir. Bununla birlikte, Random Forest gibi daha gelişmiş algoritmalara kıyasla nispeten daha düşük doğruluk oranları sunması, model seçiminde göz önünde bulundurulması gereken bir faktör olarak değerlendirilmelidir.

KNN modelinin performans değerlendirmesi, çeşitli metrikler üzerinden kapsamlı bir analiz ortaya koymaktadır. Model, 0.52'lik AUC değeri ile sınıfların ayrıştırılmasında diğer modellere kıyasla en düşük performansı sergilerken, %91'lik yüksek duyarlılık değeri ile pozitif sınıfları tanımlama konusunda dikkat çekici bir başarı göstermiştir. %72'lik kesinlik değeri ve %69'luk genel doğruluk oranı, modelin sınıflandırma görevindeki kısıtlarını vurgularken, 0.80'lik F1 skoru kesinlik ve duyarlılık metrikleri arasında makul bir denge sağladığını göstermektedir. KNN algoritmasının pozitif örnekleri tespit etmedeki yüksek başarısına rağmen, düşük AUC değeri ve artan veri boyutuyla yükselen hesaplama maliyeti, modelin genel amaçlı sınıflandırma görevlerinden ziyade, belirli kullanım senaryolarında tercih edilebileceğini ortaya koymaktadır. Bu özellikler, modelin kullanım alanlarını sınırlandıran faktörler olarak değerlendirilmekte ve uygulama seçiminde dikkatli bir analiz gerektirmektedir.

Lojistik Regresyon ve Karar Ağacı modelleri ise veri setinin genel karakteristiğini yakalamada orta düzeyde bir başarı gösterdi. Bu modeller, dengeli performansları ve yorumlanabilir yapıları sayesinde pratik uygulamalarda tercih edilebilir olmakla birlikte, daha gelişmiş algoritmalara kıyasla sınırlı bir performans sergilemektedir. Model seçiminde, uygulamanın gereksinimleri ve kısıtları göz önünde bulundurulmalıdır. Yüksek doğruluk ve ayrıştırma performansının öncelikli olduğu durumlarda Random Forest modeli öne çıkarken, hızlı işlem süresi ve kolay yorumlanabilirlik gerektiren senaryolarda Lojistik regresyon ve Karar Ağacı modelleri daha uygun alternatifler olarak değerlendirilmiştir.

#### 7.3.4.1 USG Verileri Dahil Edildiğinde Model Performansı

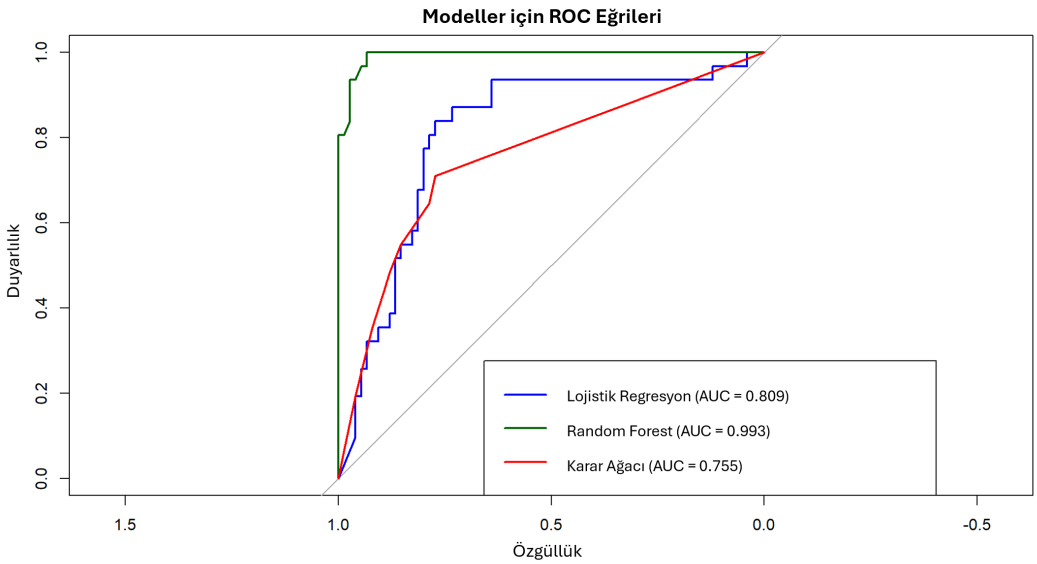
ROC (Receiver Operating Characteristic) eğrisi, Lojistik Regresyon, Random Forest ve Karar Ağacı modellerinin performanslarını USG verileriyle birlikte kıyaslamaktadır (Şekil 7.5). R programlama dilinde `plot()` ve `lines()` fonksiyonları kullanılarak oluşturulan ROC eğrilerine göre, Random Forest modeli, AUC (Area Under Curve) değeri 0.993 ile en yüksek performansı sergilemektedir. Lojistik Regresyon ve Karar Ağacı algoritmalarının AUC değerleri sırasıyla 0.858 ve 0.835 olarak kaydedilmiştir ve bu iki model, Random Forest'a kıyasla daha düşük bir performans gösterdi. ROC eğrisi, doğru pozitif oranı ile yanlış pozitif oranı arasındaki ilişkiyi gösterir ve eğrinin sol üst köşeye yakınlığı, modelin sınıflandırma başarısının bir göstergesidir. USG verilerinin modele dahil edilmesi, özellikle Random Forest modelinin başarısını artırmış ve modelin doğruluğunu önemli ölçüde geliştirdi. Bu durum, USG verilerinin prenatal tanı süreçlerinde kritik bir veri kaynağı olduğunu göstermektedir.




**Şekil 7.5** Ultrason verileri dahil edildiğinde makine öğrenmesi modellerinin ROC analizi karşılaştırması

#### 7.3.4.2 USG Verileri Hariç Tutulduğunda Model Performansı

Lojistik Regresyon, Random Forest ve Karar Ağacı modellerinin USG verisi olmadan performanslarını Şekil 7.6'te ROC eğrisi üzerinden değerlendirmektedir. Random Forest modeli, USG verisi olmaksızın da AUC değeri 0.993 ile en yüksek performansı sergiledi. Lojistik Regresyon ve Karar Ağacı modellerinin AUC değerleri ise sırasıyla 0.809 ve 0.755 olarak ölçüldü ve USG'li duruma kıyasla daha düşük performans gösterdi. Bu, USG verilerinin model performansına önemli bir katkı



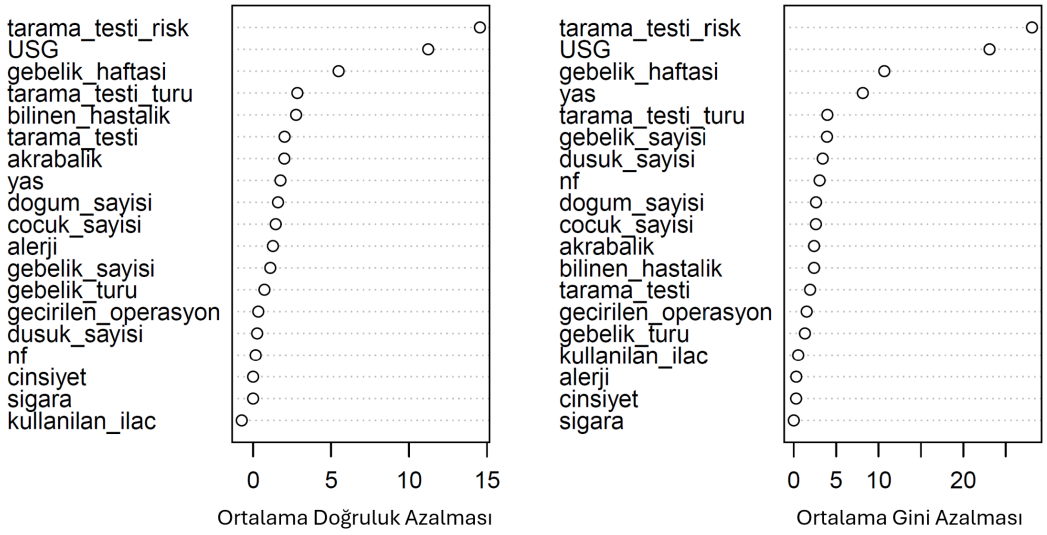

**Şekil 7.6** Ultrason verileri hariç tutulduğunda makine öğrenmesi modellerinin ROC analizi karşılaştırması

sağladığını göstermektedir. Ancak USG verisi olmadan da Random Forest'ın diğer algoritmalara göre daha iyi performans gösterdiği anlaşılmaktadır. USG verisi eksikliğine rağmen Random Forest modelinin başarısı, modelin karmaşık veri ilişkilerini daha iyi yakalayabildiğini ve diğer yöntemlere kıyasla daha güçlü bir sınıflandırma performansı sergilediğini ortaya koymaktadır.

### 7.3.5 Değişken Önem Dereceleri ve Model Performansı

Random Forest modeli için gerçekleştirilen değişken önem derecesi analizi, modelin doğruluğunu ve ayrım gücünü etkileyen faktörlerin belirlenmesi açısından değerli bilgiler sunmaktadır. Prenatal tanı sürecinde kullanılan verilerin model performansına olan katkısı, Ortalama Doğruluk Azalması (Mean Decrease Accuracy) ve Ortalama Gini Azalması (Mean Decrease Gini) olmak üzere iki farklı ölçüte göre değerlendirilmiştir. Her iki metrik de, değişkenlerin modele olan katkısını farklı perspektiflerden değerlendirerek önemli değişkenleri ön plana çıkarmıştır. Her iki ölçüte göre de tarama testi riski ve USG değişkenleri, model performansı üzerinde en yüksek etkiye sahip belirleyici değişkenler olarak saptanmıştır. Random Forest modelindeki değişkenlerin önem dereceleri Şekil 7.7'de görüldüğü üzere, Doğruluk Azalması (sol) ve Gini Saflık İndeksi (sağ) metrikleriyle gösterilmektedir. Tarama testi riski hem modelin doğruluğunu artırmada hem de Gini indeksine dayalı ayrım gücünü geliştirmede en kritik değişken olarak öne çıkmıştır. Bu değişkenin, prenatal tanıda anomali riskini öngörmede merkezi bir rol oynadığı açıkça görülmektedir.

USG ve gebelik haftası değişkenleri, sırasıyla ikinci ve üçüncü sırada yer




**Şekil 7.7** Random Forest modelinde değişkenlerin Doğruluk Düşüşü ve Gini Saflık İndeksine göre önem dereceleri

alarak model performansını olumlu yönde etkileyen diğer önemli faktörler olarak belirlenmiştir. Bu değişkenlerin, ultrasonografi bulguları ve gebelik süreciyle ilgili bilgiler sağlaması nedeniyle, özellikle klinik uygulamalarda karar verme süreçlerinde büyük bir öneme sahip olduğu görülmektedir.

Bilinen hastalık, tarama testi türü, gebelik sayısı ve yaş gibi değişkenler de model performansını orta düzeyde etkileyen faktörler arasında yer almıştır. Bu değişkenler, modelin karar mekanizmasını destekleyen ve klinik bağlamda anlamlılık taşıyan verilerdir. Buna karşılık, kullanılan ilaç, cinsiyet ve sigara gibi değişkenler, hem doğruluk hem de Gini değerleri açısından daha düşük önem derecesine sahiptir. Bu durum, bu değişkenlerin modele katkısının sınırlı olduğunu ve belki de belirli alt gruplarda daha fazla önem taşıyabileceğini düşündürmektedir. Bu değişkenler, modelin genel performansını olumsuz etkilemeyecek şekilde analize dahil edilmeyebilir veya farklı bir bağlamda yeniden değerlendirilerek kullanımları optimize edilebilir.

Sonuçlar, prenatal tanı sürecinde belirli değişkenlerin öncelikli olarak değerlendirilmesi gerektiğini ortaya koymaktadır. Özellikle tarama testi riski, USG ve gebelik haftası gibi değişkenlerin, hem klinik karar destek sistemlerinde hem de model geliştirme süreçlerinde kritik bir rol oynadığı anlaşılmaktadır. Modelin doğruluğunu ve ayrım gücünü artırmak adına bu değişkenlere odaklanılması, daha güvenilir ve etkili bir tahmin süreci sağlayabilir. Kullanılan iki farklı metrik - Doğruluk Azalması ve Gini Saflık İndeksi - bu değişkenlerin önemini farklı açılardan doğrulamıştır. Doğruluk Azalması değişkenlerin tahmin gücünü ölçerken,

Gini Saflık İndeksi veriyi ayırma kabiliyetini değerlendirmektedir. Her iki metrikte de 'tarama testi riski' ve 'USG' en önemli belirleyici değişkenler olarak öne çıkmıştır. Bu durum, söz konusu değişkenlerin hem tahmin performansı hem de veri sınıflandırması açısından güçlü göstergeler olduğunu kanıtlamaktadır. Buna ek olarak, daha az önemli değişkenlerin analizden çıkarılması ya da yeniden yapılandırılması, modelin karmaşıklığını azaltarak daha hızlı ve verimli çalışmasını sağlayabilir.

Bu bulgular, yalnızca model performansını optimize etmekle kalmayıp aynı zamanda prenatal tanı süreçlerinin iyileştirilmesine yönelik yeni stratejiler geliştirilmesine de olanak tanıyabilir. Klinik ve genetik verilerin entegre bir şekilde ele alınmasıyla, daha hassas risk tahminleri yapılabilir ve gebelik yönetiminde daha etkili yaklaşımlar benimsenebilir. Dolayısıyla, bu analizlerin sonuçları, yalnızca bir model değerlendirmesi değil, aynı zamanda klinik uygulamalara yönelik pratik bir rehber olarak da değerlendirilebilir.

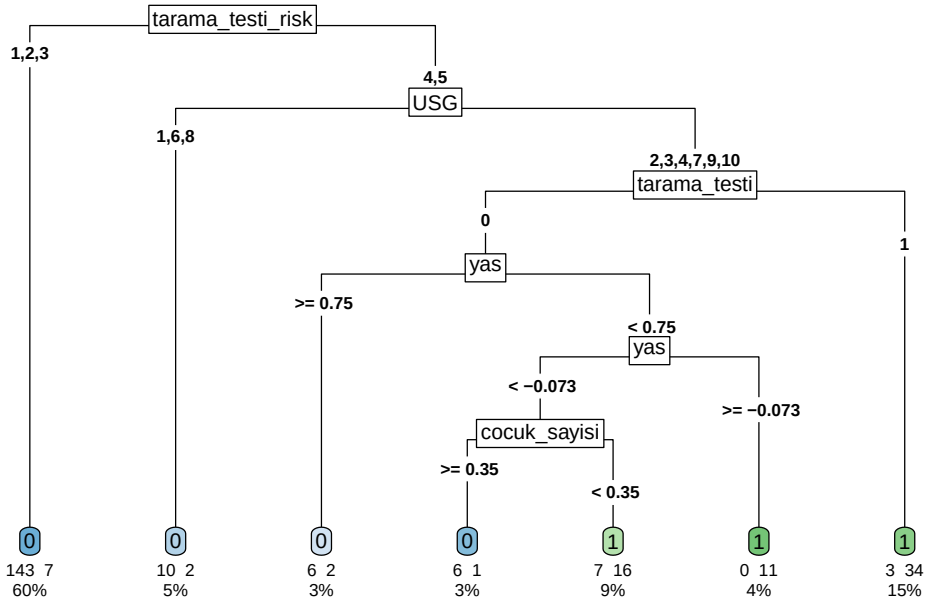
### 7.3.6 Karar Ağacı Modelinin Analizi

Karar Ağacı modeli, düşük hesaplama maliyeti ve kolay yorumlanabilir yapısı ile öne çıkmaktadır. Bununla birlikte, model özellikle büyük veri setlerinde gözlemlenen aşırı öğrenme eğilimi ve Random Forest modeline kıyasla daha düşük AUC değeri gibi sınırlayıcı faktörlere sahiptir. Çalışmamızda, modelin her bir özelliği trizomi riski ile ilişkilendirilerek dallara ayrılmış ve model derinliği, performans ile aşırı öğrenme riski arasındaki denge gözetilerek optimize edilmiştir.

Karar ağacı modeli, hızlı işlem süresi ve yüksek açıklanabilirlik özelliklerini korurken, aşırı öğrenme riski düşük olacak şekilde tasarlanmıştır. Modelin performansını artırmak için, çapraz doğrulama yöntemleri kullanılarak karmaşıklık parametresi (CP) optimize edilmiştir. Bunun sonucunda elde edilen ağaç, yalnızca en anlamlı değişkenleri kullanarak dallanmalar yapmış ve bu sayede klinik uygulamalar için yoruma açıklık sağlamıştır. Nihai değerlendirmede, karar ağacı modelinin sınıflandırma performansının (AUC = 0.83) Random Forest (AUC = 0.99) gibi daha karmaşık modellerin gerisinde kaldığı gözlemlenmiştir.

Bu sonuçlar, model seçiminde veri seti özellikleri, uygulama gereksinimleri ve modelin genelleştirme yeteneğinin dikkatle değerlendirilmesi gerektiğini vurgulamaktadır. Ayrıca, modelin sınıflandırma görevindeki ayrıştırma kabiliyetinin diğer modellere göre sınırlı olduğunu göstermektedir. Modelin 0.88'lük F1 skoru, kesinlik ve duyarlılık metrikleri arasında başarılı bir denge yakaladığını işaret etmektedir. Bu değer, modelin hem doğru pozitifleri tespit etme hem de

yanlış pozitifleri minimize etme konusunda etkili olduğunu ortaya koymaktadır. %82'lık genel doğruluk oranı, Lojistik Regresyon modeline yakın bir performans sergilediğini ve veri setindeki örneklerin önemli bir bölümünü doğru şekilde sınıflandırabildiğini göstermektedir. Duyarlılık metriğinde elde edilen %92'lik başarı oranı, modelin pozitif sınıfları tanımlama konusundaki yetkinliğini vurgularken, %84'lik kesinlik değeri, pozitif sınıf tahminlerindeki güvenilirliği desteklemektedir. Bu metrikler, modelin genel sınıflandırma performansının tatmin edici düzeyde olduğunu göstermektedir.




**Şekil 7.8** QF-PCR sonuçlarının tahmininde kullanılan Karar Ağacı modeli

Karar ağacı, modelin çalışmada kullanılan verideki temel desenleri öğrenme kapasitesini göstermektedir. Şekil 7.8, karar ağacının dallanma yapısını ve sınıflandırma süreçlerinde kullanılan değişkenleri öne çıkarmaktadır. QF-PCR sonuçlarının tahmininde kullanılan Karar Ağacı modeli, çalışma verisi üzerinde bireyleri tarama testi riski, USG, tarama testi ve yaş değişkenlerine göre alt gruplara ayırarak sınıflandırmıştır. Mavi ve yeşil düğümler sırasıyla normal (sınıf 0) ve anomali (sınıf 1) olan QF-PCR sonuçlarını temsil etmekte, renk tonları ise sınıf dağılımını göstermektedir. Bu süreçte kullanılan yüzdelik oranlar, modelin belirli gruplar içindeki sınıflandırma performansını yansıtırken, eşik değerleri değişkenlerin önem sırasını belirlemektedir.

Karar ağacında kullanılan yüzdeler (%60, %15 gibi), modelin belirli gruplar içinde yaptığı sınıflandırma dağılımını ve bu grupların genel veri içerisindeki oranlarını temsil ederken; 0.75, -0.073 ve 0.35 gibi eşik değerleri ise sırasıyla yaş, çocuk sayısı

ve USG değişkenlerinin sınıflandırma sürecindeki belirleyici rolünü göstermektedir. Bu eşik değerler, modelin her bir yaprak düğümde hangi değişkenlere öncelik verdiğini ve sınıflandırılan verinin genel dağılımını anlamamıza yardımcı olur. Böylece, sınıflandırılan her grubun çalışma verisindeki yüzdesi ve bu grupların klinik sonuçlarla olan ilişkisi daha net bir şekilde ortaya konulabilir.

Karar ağacının kök düğümünde tarama testi riski değişkeni yer almakta ve bu değişken, sınıflandırma sürecindeki temel ayrım noktası olarak değerlendirilmektedir. Ardından USG değişkeni, QF-PCR sonuçlarının daha ayrıntılı ayrımında kritik bir rol oynamaktadır. Örneğin, "USG = 1,6,8" olarak değerlendirilen bireylerin %5'i "0" sınıfına (normal QF-PCR sonuçlarına) atanmıştır. Bu durum düşük USG sonuçlarının, genellikle normal gebelik sonuçlarıyla ilişkili olduğunu göstermektedir.

USG = 2, 3, 4, 7, 9 ve 10 durumunda ise, tarama testi değişkeni devreye girmekte ve bireylerin %15'i "1" sınıfına (anomali olan QF-PCR sonuçlarına) atanmıştır. Bu bulgu, tarama testi sonuçlarının riskli gebeliklerin belirlenmesinde önemli bir rol oynadığını ortaya koymaktadır. USG'de daha yüksek riskli gruba dahil edilmiş ve bu grupta anomali olan QF-PCR sonuçlarının daha sık gözlemlendiği tespit edilmiştir.

Tarama testi sonucu = 0 olduğunda, yaş ve çocuk sayısı gibi değişkenler ayrımı daha ileri düzeyde belirlemektedir. Örneğin, yaş $\geq 0.75$ olarak değerlendirilen bireylerin %3'ü "0" sınıfına atanmış, bu durum ileri yaş grubu bireylerin genellikle normal sonuçlarla ilişkilendirildiğini göstermiştir. Buna karşılık, yaş $< 0.75$ olan bireylerde daha fazla detay ayrımı yapılmış ve yaş $< -0.073$ olarak değerlendirilen bireylerin %9'unun "1" sınıfına (anomali olan QF-PCR sonuçlarına) atanması, bu grupta olumsuz gebelik sonuçlarının daha yaygın olduğunu göstermektedir.

Çocuk sayısı $\geq 0.35$ olan bireylerin tamamı "0" sınıfına (normal sonuçlar) atanırken, çocuk sayısı $< 0.35$ olan bireylerin %9'u "1" sınıfına atanmıştır. Bu durum, düşük çocuk sayısının olumsuz gebelik sonuçlarıyla ilişkilendirilebileceğini göstermektedir.

Karar ağacında dallar boyunca kullanılan değişkenler ve eşik değerleri, gebelik sonuçlarının öngörülmesi ve riskli grupların belirlenmesi için anlamlı bir sınıflandırma süreci sunmaktadır. Özellikle tarama testi sonucu = 1 olan bireylerin %15'inin "1" sınıfına atanması, bu testin gebelik risk değerlendirmesindeki kritik önemini doğrulamaktadır.

Sonuç olarak, bu karar ağacı modeli USG, yaş, çocuk sayısı gibi değişkenlerin QF-PCR sonuçları üzerindeki etkilerini anlamak ve risk gruplarını belirlemek için etkili bir araç sunmaktadır. Bu bulgular, klinik uygulamalarda karar verme süreçlerini desteklemek ve erken müdahale stratejilerini geliştirmek için rehberlik edebilir.

Karar ağacında kullanılan renkler, düğümlerdeki sınıf oranlarını ve sınıflandırma sonuçlarını görselleştirmek için kullanılmıştır:

- Mavi Düğümler: Sınıf 0 (normal QF-PCR sonuçları) ağırlıklı düğümleri temsil eder.
- Yeşil Düğümler: Sınıf 1 (anomali olan QF-PCR sonuçları) ağırlıklı düğümleri ifade eder.
- Renk Tonları: Renk tonlarının yoğunluğu, düğümdeki sınıf ağırlığını gösterir. Daha koyu renkler, bir sınıfın daha baskın olduğunu belirtir.

Sonuç olarak Karar Ağacı modeli, klinik uygulamalarda hızlı karar verme süreçlerini desteklemek için uygun bir araç sunarken, daha karmaşık modellerle kıyaslandığında performansın bazı durumlarda sınırlı kalabileceğini göstermiştir. Her ne kadar hızlı işlem süresi ve yüksek açıklanabilirlik özelliklerini korusa da, nihai değerlendirmede Karar Ağacı modelinin diğer yöntemlere (Lojistik Regresyon ve Random Forest'a) kıyasla daha düşük performans sergilediği gözlemlenmiştir. Bu sonuçlar, model seçiminde veri seti özellikleri ve uygulama gereksinimlerinin dikkatli bir şekilde değerlendirilmesi gerektiğini vurgulamaktadır.

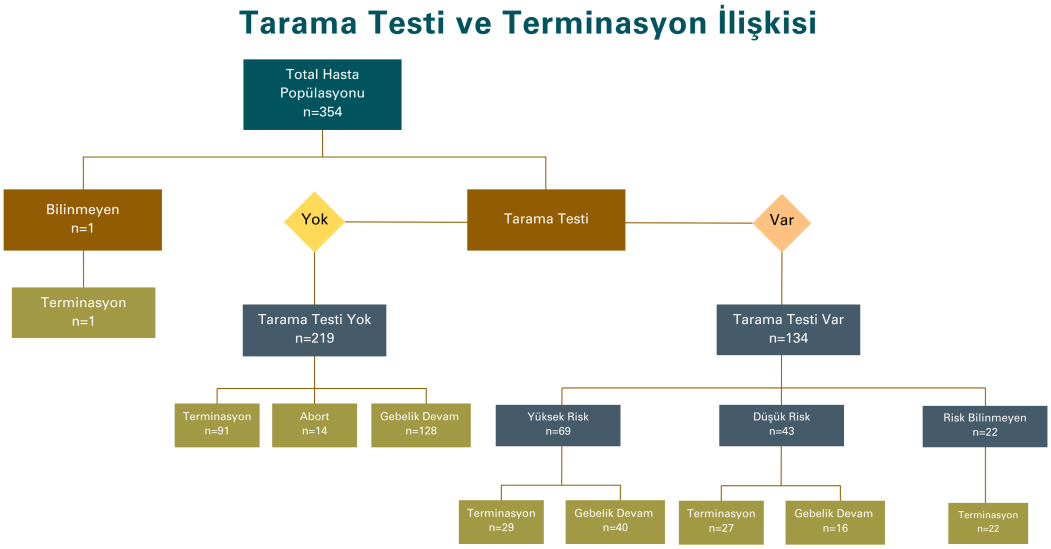
## 7.4 Prenatal Tarama Testleri ve Terminasyon Kararlarının Değerlendirilmesi

Tarama testi yaptıran ve yaptırmayan grupların risk değerlendirmesi ve terminasyon arasındaki ilişki Şekil 7.9'da gösterilmiştir. Bu şekilde toplam 354 gebelikte prenatal tarama testi uygulamasının (n=134) ve uygulanmamasının (n=219) gebelik sonuçlarına etkisi incelenmiştir. Tarama testi uygulanan grupta risk düzeylerine göre terminasyon ve gebelik devamı oranları gösterilmiştir.

Genel popülasyon dağılımında prenatal tanı için başvuran hastaların %37.8'i (134/354) tarama testi yaptırırken, çoğunluk (%61.9, 219/354) tarama testi yaptırmamıştır. Popülasyonun %0.3'ünde tarama durumu bilgisi mevcut değildir.

Beklenmedik şekilde, tarama testi yaptıran grupta düşük riskli hastaların terminasyon oranı (%62.8), yüksek riskli gruba (%42) göre daha yüksek bulunmuştur. Bu bulgu, terminasyon kararlarının sadece risk değerlendirmesine bağlı olmadığını göstermektedir. Bu durum, hasta tercihlerinin sosyoekonomik, psikolojik ve kültürel faktörlerden etkilenebileceğini düşündürmektedir.

Önemli bulgular şunlardır:




**Şekil 7.9** Prenatal tarama test sonuçlarının risk kategorilerine göre terminasyon ve gebelik sonucu tahmini üzerindeki etkisinin analizi

- Risk değeri bilinmeyen grupta terminasyon oranı %100'e ulaşmıştır. Bu grubun örneklem büyüklüğü oldukça küçük olup, istatistiksel genellemeler yapmak için yetersiz olabilir. Ancak bu bulgu, klinik kararların yalnızca risk değerlendirmesine bağlı olmadığını ortaya koymaktadır.
- Tarama testi olmayan grupta %41.5 oranında terminasyon kararı alınmıştır.
- Terminasyon kararlarında hasta tercihleri ve diğer klinik bulgular belirleyici rol oynamaktadır.

Makine öğrenmesi analizleri, terminasyon kararının çok faktörlü yapısını doğrulanmıştır. Bu çalışmada lojistik regresyon ve karar ağaçları gibi makine öğrenmesi yöntemleri kullanılmış; USG bulguları, QF-PCR sonuçları ve hasta yaşı gibi değişkenler modele dahil edilmiştir. QF-PCR sonuçları ve USG bulguları gibi objektif verilerin modele entegrasyonu, sistemin güvenilirliğini artırmıştır. Yüksek duyarlılık değerleri, riskli vakaların başarılı tespitini gösterirken; Kesinlik değerleri, yanlış pozitif oranının kabul edilebilir seviyelerde olduğunu ortaya koymuştur.

Bu sonuçlar, terminasyon kararlarının tek bir faktöre bağlı olmadığını, ultrason bulguları, genetik test sonuçları ve hasta tercihlerinin bütüncül değerlendirilmesi

gerektiğini göstermektedir. Ayrıca, multidisipliner ekiplerin ve hasta danışmanlığı süreçlerinin bireyselleştirilmesi, karar sürecindeki hataların önüne geçilmesinde önemli bir rol oynayabilir. Makine öğrenmesi modelleri, bu karmaşık karar sürecinin analizinde ve risk faktörlerinin ağırlıklandırılmasında değerli bilgiler sunmaktadır. Risk seviyesinin tek başına belirleyici olmadığı görülmüş, hasta tercihlerinin ve diğer klinik bulguların karar sürecindeki etkisi makine öğrenmesi sonuçlarıyla desteklenmiştir. Bu durum, klinisyenlerin terminasyon kararları için daha esnek bir yaklaşım benimsemesi gerektiğini ve bireysel hasta ihtiyaçlarının göz önünde bulundurulmasının önemini ortaya koymaktadır. Bu bulgular, prenatal tanı sürecinde bütüncül bir yaklaşımın önemini vurgulamaktadır.

## 7.5 Prenatal Tanı Süreçlerinde Makine Öğrenmesi Modellerinin Entegrasyonu ve Trizomi Risk Tahmini

Prenatal tanı süreçlerinde makine öğrenmesi modellerinin kullanımı ve trizomi riskinin erken dönemde tahmin edilmesi konusunda yürütülen bu çalışmada, gebelik yönetiminde veri odaklı yaklaşımların etkinliği araştırılmıştır. Bu çalışma, yalnızca mevcut risk değerlendirme yöntemlerini optimize etmekle kalmayıp, yeni algoritmaların entegrasyonu ile prenatal tanı süreçlerini hızlandırmayı ve klinik karar destek sistemlerini güçlendirmeyi hedeflemiştir.

Makine öğrenmesi modelleri arasında özellikle Random Forest modelinin kullanımı, çok sayıda parametrenin değerlendirilmesinde ve karmaşık ilişkilerin analizinde başarılı sonuçlar verdi. Doğrusal olmayan ilişkilerin varlığı ve kategorik verilerin yoğunluğu nedeniyle Random Forest'ın tercih edilmesi, özellikle ensemble yöntemlerin karmaşık medikal verilerdeki üstün performansıyla desteklenmiştir. Random Forest, aşırı öğrenme eğiliminden korunma mekanizması sayesinde genelleme yeteneği yüksek bir model olarak öne çıkmıştır. Random Forest modeli, %99 AUC ve %87 F1 skoru ile diğer modellerden daha başarılı performans göstermiştir. Örneğin, lojistik regresyon modeli AUC = %86, karar ağacı modeli ise AUC = %83 performans göstermiştir. Bu karşılaştırma, Random Forest'ın doğruluk ve genelleme açısından üstünlüğünü ortaya koymaktadır.

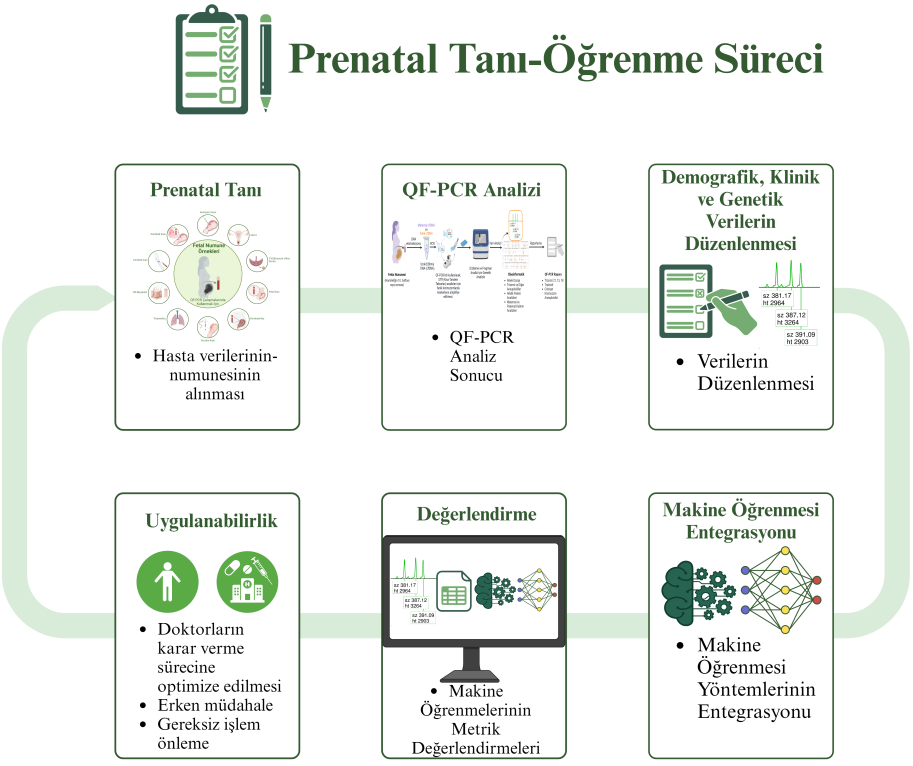
Çalışmada geliştirilen sistemin klinik uygulamaya katkıları şu şekilde özetlenebilir:

**Erken Teşhis:** Makine öğrenmesi algoritmaları, klasik yöntemlerle kıyaslandığında büyük veri kümeleri üzerinde daha hızlı tahmin yapabilmekte ve erken müdahale sürecini hızlandırmaktadır.

**Doğruluk Oranı:** Yüksek doğruluk oranları, manuel yöntemlere kıyasla daha

güvenilir tahminler sunmaktadır.

**Klinik Karar Destek:** Model entegrasyonu, gereksiz tıbbi prosedürleri önleyerek maternal-fetal sağlığın optimizasyonuna katkı sağlamaktadır. Özellikle invaziv prosedürlerden kaçınmak isteyen hastalar için model, güvenilir ve invaziv olmayan bir alternatif sunmaktadır.




**Şekil 7.10** Prenatal tanıda makine öğrenmesi entegrasyonunun sistematik iş akış şeması, Biorender ile oluşturuldu.

Şekil 7.10'de gösterilen prenatal tanı makine öğrenmesi entegrasyonunun sistematik iş akış süreçlerinin şeması, prenatal tanı sürecinin veri toplama aşamasından makine öğrenmesi uygulamasına kadar olan iş akışını göstermektedir. QF-PCR analizleri, klinik verilerin düzenlenmesi ve makine öğrenmesi entegrasyonu ile prenatal tanı sürecinin optimize edilmesi hedeflenmektedir. Bu süreç, hastalardan alınan klinik ve genetik verilerin analizi ile başlamakta, QF-PCR analizi ile genetik bozuklukların hızlı tespiti sağlanmaktadır. Demografik, klinik ve genetik veriler düzenlenerek analiz için hazır hale getirilmekte ve makine öğrenmesi modelleriyle işlenmektedir. Karar ağacı görselleri ve değişken önem dereceleri grafikleri, sınıflandırma sürecindeki kritik faktörleri ortaya koyarken, ROC eğrileri özellikle USG verilerinin modele katkısını net bir şekilde göstermiştir. Bu süreçte veriler R programlama dili kullanılarak analiz edilmiş ve sonuçlar görselleştirilmiştir.

Özellikle invaziv yöntemlerden kaçınan hastalar ve invaziv olmayan risk belirleme yöntemlerine ihtiyaç duyan hekimler için, geliştirilen model önemli bir alternatif sunmaktadır. Bununla birlikte, modelin küçük veri setleri üzerindeki performansı ve genel uygulanabilirliği, daha geniş veri setlerinde test edilerek doğrulanmalıdır. Parametrelerin disiplinli kullanımı, hem sağlık sistemi hem de hasta için süreç içerisinde daha efektif aksiyon alınmasını sağlamaktadır. Bu çalışma ile geliştirilen yaklaşım, trizomi risk tahmininde klinik süreçleri daha etkili hale getirerek hem akademik hem de klinik alanda yenilikçi katkılar sağlamaktadır.

# KAYNAKÇA


[1] K. Boddupally E. Rani Thuraka, "Artificial intelligence for prenatal chromosome analysis," *Clinica Chimica Acta*, c. 552, s. 117 669, Ocak 2024. doı: 10.1016/j.cca.2023.117669.

[2] M. Shaban, S. Mollazadeh, S. Eslami, F. Tara, S. Sharif F. E. Arghavanian, "Prediction of chromosomal abnormalities in the screening of the first trimester of pregnancy using machine learning methods: a study protocol," *Reproductive Health*, c. 21, no. 1, Tem. 2024. doı: 10.1186/s12978-024-01839-5.

[3] M. Hulten, S. Dhanjal B. Pertl, "Rapid and simple prenatal diagnosis of common chromosome disorders: advantages and disadvantages of the molecular methods FISH and QF-PCR," *Reproduction*, c. 126, no. 3, ss. 279–297, Eyl. 2003. doı: 10.1530/rep.0.1260279.

[4] U. Nicolini, F. Lalatta, F. Natacci, C. Curcio T.-H. Bui, "The introduction of QF-PCR in prenatal diagnosis of fetal aneuploidies: time for reconsideration," *Human Reproduction Update*, c. 10, no. 6, ss. 541–548, Kas. 2004. doı: 10.1093/humupd/dmh046.

[5] S. Langlois ve diğ., "Use of a DNA Method, QF-PCR, in the Prenatal Diagnosis of Fetal Aneuploidies," *Journal of Obstetrics and Gynaecology Canada*, c. 33, no. 9, ss. 955–960, Eyl. 2011. doı: 10.1016/s1701-2163(16)35022-8.

[6] K. N. Sreelakshmi, "Medical Genetics for Practicing Obstetrician," *The Journal of Obstetrics and Gynecology of India*, c. 70, no. 1, ss. 6–11, Ağus. 2019. doı: 10.1007/s13224-019-01257-9.

[7] T. Sládeček ve diğ., "Combination of expert guidelines-based and machine learning-based approaches leads to superior accuracy of automated prediction of clinical effect of copy number variations," *Scientific Reports*, c. 13, no. 1, Haz. 2023. doı: 10.1038/s41598-023-37352-1.

[8] F. Brioude ve diğ., "Clinical and molecular diagnosis, screening and management of Beckwith–Wiedemann syndrome: an international consensus statement," *Nature Reviews Endocrinology*, c. 14, no. 4, ss. 229–249, Ocak 2018. doı: 10.1038/nrendo.2017.166.

[9] I. Bedei, A. Wolter, A. Weber, F. Signore R. Axt-Fliedner, "Chances and Challenges of New Genetic Screening Technologies (NIPT) in Prenatal Medicine from a Clinical Perspective: A Narrative Review," *Genes*, c. 12, no. 4, s. 501, Mar. 2021. doı: 10.3390/genes12040501.

[10] S. Eren Keskin ve diğ., "Cytogenetic evaluation of 661 prenatal samples," *Cukurova Medical Journal*, c. 49, no. 2, ss. 248–259, Haz. 2024. doı: 10.17826/cumj.1380467.

[11] O. Pös, J. Budiš T. Szemes, "Recent trends in prenatal genetic screening and testing," *F1000Research*, c. 8, s. 764, May. 2019. doı: 10.12688/f1000research.16837.1.

[12] M. De Rycke V. Berckmoes, "Preimplantation Genetic Testing for Monogenic Disorders," *Genes*, c. 11, no. 8, s. 871, Tem. 2020. doı: 10.3390/genes11080871.

[13] K. Mann C. M. Ogilvie, "QF-PCR: application, overview and review of the literature," *Prenatal Diagnosis*, c. 32, no. 4, ss. 309–314, Mar. 2012. doı: 10.1002/pd.2945.

[14] A. D. Adams, S. Stover M. W. Rac, "Omphalocele—What should we tell the prospective parents?" *Prenatal Diagnosis*, c. 41, no. 4, ss. 486–496, Şub. 2021. doı: 10.1002/pd.5886.

[15] V. Wolters ve diğ., "Management of pregnancy in women with cancer," *International Journal of Gynecologic Cancer*, c. 31, no. 3, ss. 314–322, Mar. 2021. doı: 10.1136/ijgc-2020-001776.

[16] L. Rickman, H. Fiegler, N. Carter M. Bobrow, "Prenatal Diagnosis by Array-CGH," *European Journal of Medical Genetics*, c. 48, no. 3, ss. 232–240, Tem. 2005. doı: 10.1016/j.ejmg.2005.03.003.

[17] K. O. Kagan, J. Sonek P. Kozlowski, "Antenatal screening for chromosomal abnormalities," *Archives of Gynecology and Obstetrics*, c. 305, no. 4, ss. 825–835, Mar. 2022. doı: 10.1007/s00404-022-06477-5.

[18] C. Xu, J. Peng, Y. Zhang, S. Liang D. Wang, "Detection of partial deletion and mosaicism using quantitative fluorescent polymerase chain reaction: Case reports and a review of the literature," *Journal of Clinical Laboratory Analysis*, c. 36, no. 8, Haz. 2022. doı: 10.1002/jcla.24574.

[19] M. Agathokleous, P. Chaveeva, L. C. Y. Poon, P. Kosinski K. H. Nicolaides, "Meta-analysis of second-trimester markers for trisomy 21," *Ultrasound in Obstetrics and Gynecology*, c. 41, no. 3, ss. 247–261, Ocak 2013. doı: 10.1002/uog.12364.

[20] T. M. MacDonald, S. P. Walker, N. J. Hannan, S. Tong T. J. Kaitu'u-Lino, "Clinical tools and biomarkers to predict preeclampsia," *eBioMedicine*, c. 75, s. 103 780, Ocak 2022. doı: 10.1016/j.ebiom.2021.103780.

[21] M. Abedalthagafi, S. Bawazeer, R. I. Fawaz, A. M. Heritage, N. M. Alajaji E. Faqeih, "Non-invasive prenatal testing: a revolutionary journey in prenatal testing," *Frontiers in Medicine*, c. 10, Kas. 2023. doı: 10.3389/fmed.2023.1265090.

[22] S. Paul, A. Nahar, M. Bhagawati A. J. Kunwar, "A Review on Recent Advances of Cerebral Palsy," *Oxidative Medicine and Cellular Longevity*, c. 2022, G. Kumar, ed., ss. 1–20, Tem. 2022. doı: 10.1155/2022/2622310.

[23] D. Levenson, "Cell-free Fetal DNA Tests Appropriately Used by Geneticists, Often Misunderstood by Patients: Survey finds many patients incorrectly think prenatal screening test is diagnostic," *American Journal of Medical Genetics Part A*, c. 164, no. 4, Mar. 2014. doı: 10.1002/ajmg.a.36536.

[24] J. M. H. Horsting, S. R. Dlouhy, K. Hanson, K. Quaid, S. Bai K. A. Hines, "Genetic Counselors' Experience with Cell-Free Fetal DNA Testing as a Prenatal Screening Option for Aneuploidy," *Journal of Genetic Counseling*, c. 23, no. 3, ss. 377–400, Ara. 2013. doı: 10.1007/s10897-013-9673-4.

[25] C. Meller ve diğ., "Congenital heart disease, prenatal diagnosis and management," *Archivos argentinos de pediatria*, c. 118, no. 2, e149–e161, 2020. doı: 10.5546/aap.2020.eng.e149.

[26] L. Zaninović, M. Bašković, D. Ježek A. Katušić Bojanac, "Accuracy of Non-Invasive Prenatal Testing for Duchenne Muscular Dystrophy in Families at Risk: A Systematic Review," *Diagnostics*, c. 13, no. 2, s. 183, Ocak 2023. doı: 10.3390/diagnostics13020183.

[27] P. Benn A. Rebarber, "Non-invasive prenatal testing in the management of twin pregnancies," *Prenatal Diagnosis*, c. 41, no. 10, ss. 1233–1240, Haz. 2021. doı: 10.1002/pd.5989.

[28] M. Adinolfi, B. Pertl J. Sherlock, "Rapid detection of aneuploidies by microsatellite and the quantitative fluorescent polymerase chain reaction," *Prenatal Diagnosis*, c. 17, no. 13, ss. 1299–1311, 1997.

[29] A. Asim, A. Kumar, S. Muthuswamy, S. Jain S. Agarwal, ""Down syndrome: an insight of the disease"," *Journal of Biomedical Science*, c. 22, no. 1, Haz. 2015. doı: 10.1186/s12929-015-0138-y.

[30] M. Avci, A. Arkut, N. Bilgic H. Sutcu, "The Evaluation of Invasive Prenatal Diagnostic Tests in North Cyprus: A Retrospective Study," *Nigerian Journal of Clinical Practice*, c. 27, no. 2, ss. 188–193, Şub. 2024. doı: 10.4103/njcp.njcp_540_23.

[31] S. Sharda S. R. Phadke, "Uptake of invasive prenatal diagnostic tests in women after detection of soft markers for chromosomal abnormality on ultrasonographic evaluation," *Journal of Perinatology*, c. 27, no. 9, ss. 550–555, Tem. 2007. doı: 10.1038/sj.jp.7211787.

[32] I. Herraiz ve diğ., "Impact of cell-free fetal DNA on invasive prenatal diagnostic tests in a real-world public setting," *Journal of Perinatal Medicine*, c. 47, no. 5, ss. 547–552, Mar. 2019. doı: 10.1515/jpm-2018-0410.

[33] C. Cadrin M. S. Golbus, "Fetal Tissue Sampling-Indications, Techniques, Complications, and Experience With Sampling of Fetal Skin, Liver, and Muscle," *The Western Journal of Medicine*, c. 159, no. 3, ss. 269–272, 1993.

[34] N. L. Vora L. Hui, "Next-generation sequencing and prenatal 'omics: advanced diagnostics and new insights into human development," *Genetics in Medicine*, c. 20, no. 8, ss. 791–799, Ağus. 2018. doı: 10.1038/s41436-018-0087-4.

[35] N. C. Peddi, C. Avanthika, S. Vuppalapati, R. Balasubramanian, J. Kaur C. D. N, "A Review of Cordocentesis: Percutaneous Umbilical Cord Blood Sampling," *Cureus*, Tem. 2021. doı: 10.7759/cureus.16423.

[36] C. J. Mercaldi S. F. Lanes, "Ultrasound Guidance Decreases Complications and Improves the Cost of Care Among Patients Undergoing Thoracentesis and Paracentesis," *Chest*, c. 143, no. 2, ss. 532–538, Şub. 2013. doı: 10.1378/chest.12-0447.

[37] Y. Raymond ve diğ., "Placental, maternal, fetal, and technical origins of false-positive cell-free DNA screening results," *American Journal of Obstetrics and Gynecology*, c. 230, no. 4, ss. 381–389, Nis. 2024. doı: 10.1016/j.ajog.2023.11.1240.

[38] W. C. Leung, E. T. Lau, T. T. Lao M. H. Tang, "Rapid aneuploidy screening (FISH or QF-PCR): the changing scene in prenatal diagnosis?" *Expert Review of Molecular Diagnostics*, c. 4, no. 3, ss. 333–337, May. 2004. doı: 10.1586/14737159.4.3.333.

[39] G. Biopharma, *Aneusure QF-PCR Kit Product User Manual*, 2024. erişim adresi: https://genetek-biopharma.com/wp-content/uploads/2024/09/aneusure_product-user-manual.pdf.

[40] C.-P. Chen ve diğ., "Mosaic trisomy 12 at amniocentesis: Prenatal diagnosis and molecular genetic analysis," *Taiwanese Journal of Obstetrics and Gynecology*, c. 52, no. 1, ss. 97–105, Mar. 2013. doı: 10.1016/j.tjog.2013.01.012.

[41] G. Biopharma, *A Profile of Patau Syndrome (13) Using Aneusure QF-PCR Kit*, 2024. erişim adresi: https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-patau-syndrome-13-using-aneusure-qf-pcr-kit.pdf.

[42] G. Biopharma, *A Profile of Edward Syndrome (18) Using Aneusure QF-PCR Kit*, 2024. erişim adresi: https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-edward-syndrome-18-using-aneusure-qf-pcr-kit.pdf.

[43] G. Biopharma, *A Profile of Down Syndrome (21) Using Aneusure QF-PCR Kit*, 2024. erişim adresi: https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-down-syndrome-21-using-aneusure-qf-pcr-kit.pdf.

[44] G. Biopharma, *A Profile of Turner Syndrome (XO) Using Aneusure QF-PCR Kit*, https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-turner-syndrome-xo-using-aneusure-qf-pcr-kit.pdf, 2024.

[45] M. Badeau ve diğ., "Genomics-based non-invasive prenatal testing for detection of fetal chromosomal aneuploidy in pregnant women," *Cochrane Database of Systematic Reviews*, c. 2017, no. 11, Kas. 2017. doı: 10.1002/14651858.cd011767.pub2.

[46] N. Tartaglia ve diğ., "Early neurodevelopmental and medical profile in children with sex chromosome trisomies: Background for the prospective eXtraordinarY babies study to identify early risk factors and targets for intervention," *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, c. 184, no. 2, ss. 428–443, Haz. 2020. doı: 10.1002/ajmg.c.31807.

[47] G. Biopharma, *A Profile of Klinefelter Syndrome (XXY) Using Aneusure QF-PCR Kit*, https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-klinefelter-syndrome-xxy-using-aneusure-qf-pcr-kit.pdf, 2024.

[48] T. Green, S. Flash A. L. Reiss, "Sex differences in psychiatric disorders: what we can learn from sex chromosome aneuploidies," *Neuropsychopharmacology*, c. 44, no. 1, ss. 9–21, Tem. 2018. doı: 10.1038/s41386-018-0153-2.

[49] G. Biopharma, *A Profile of Triple X Syndrome Using Aneusure QF-PCR Kit*, https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-triple-x-syndrome-using-aneusure-qf-pcr-kit.pdf, 2024.

[50] A. A. Vieira, I. Almada-Correia, J. Inácio, P. Costa-Reis S. T. da Rocha, "Female-bias in systemic lupus erythematosus: How much is the X chromosome to blame?" *Biology of Sex Differences*, c. 15, no. 1, Ekim 2024. doı: 10.1186/s13293-024-00650-y.

[51] D. Massalska, J. Bijok, A. Ilnicka, G. Jakiel T. Roszkowski, "Triploidy - variability of sonographic phenotypes: Ultrasonographic phenotypes of fetal triploidy," *Prenatal Diagnosis*, c. 37, no. 8, ss. 774–780, Tem. 2017. doı: 10.1002/pd.5080.

[52] L. Benítez ve diğ., "The Contribution of QF-PCR and Pathology Studies in the Diagnosis of Diandric Triploidy/Partial Mole," *Diagnostics*, c. 11, no. 10, s. 1811, Eyl. 2021. doı: 10.3390/diagnostics11101811.

[53] G. Biopharma, *A Profile of 69,XXX Triploid Using Aneusure QF-PCR Kit*, https://genetek-biopharma.com/wp-content/uploads/2024/09/a-profile-of-69-xxxtriploid-using-aneusure-qf-pcr-kit.pdf, 2024.

[54] D. E. McFadden, "Dispermy–origin of diandric triploidy: Brief Communication," *Human Reproduction*, c. 17, no. 12, ss. 3037–3038, Ara. 2002. doı: 10.1093/humrep/17.12.3037.

[55] D. Massalska ve diğ., "Distribution of diandric and digynic triploidy depending on gestational age," *Journal of Assisted Reproduction and Genetics*, c. 38, no. 9, ss. 2391–2395, May. 2021. doı: 10.1007/s10815-021-02202-4.

[56] N. Shreeve ve diğ., "Incremental yield of whole-genome sequencing over chromosomal microarray analysis and exome sequencing for congenital anomalies in prenatal period and infancy: systematic review and meta-analysis," *Ultrasound in Obstetrics and Gynecology*, c. 63, no. 1, ss. 15–23, Ocak 2024. doı: 10.1002/uog.27491.

[57] S. Liu ve diğ., "Non-invasive prenatal test findings in 41, 819 pregnant women: results from a clinical laboratory in southern China," *Archives of Gynecology and Obstetrics*, c. 308, no. 3, ss. 787–795, Ocak 2023. doı: 10.1007/s00404-022-06908-3.

[58] C.-P. Chen ve diğ., "Prenatal diagnosis of mosaic trisomy 8: Clinical report and literature review," *Taiwanese Journal of Obstetrics and Gynecology*, c. 50, no. 3, ss. 331–338, Eyl. 2011. doı: 10.1016/j.tjog.2011.07.013.

[59] B. H. Faas, V. Cirigliano T.-H. Bui, "Rapid methods for targeted prenatal diagnosis of common chromosome aneuploidies," *Seminars in Fetal and Neonatal Medicine*, c. 16, no. 2, ss. 81–87, Nis. 2011. doı: 10.1016/j.siny.2011.01.003.

[60] M.-L. Sun ve diğ., "Prenatal diagnosis and molecular cytogenetic characterization of a small supernumerary marker chromosome (sSMC) inherited from her mosaic sSMC(15) mother and a literature review," *Taiwanese Journal of Obstetrics and Gynecology*, c. 59, no. 6, ss. 963–967, Kas. 2020. doı: 10.1016/j.tjog.2020.09.030.

[61] J. D. Delhanty, S. B. SenGupta H. Ghevaria, "How common is germinal mosaicism that leads to premeiotic aneuploidy in the female?" *Journal of Assisted Reproduction and Genetics*, c. 36, no. 12, ss. 2403–2418, Kas. 2019. doı: 10.1007/s10815-019-01596-6.

[62] H. Zhang ve diğ., "Prenatal Diagnosis and Molecular Cytogenetic Characterization of Copy Number Variations on 4p15.2p16.3, Xp22.31, and 12p11.1q11 in a Fetus with Ultrasound Anomalies: A Case Report and Literature Review," *BioMed Research International*, c. 2020, ss. 1–10, May. 2020. doı: 10.1155/2020/1761738.

[63] N. Zeinali, N. Youn, A. Albashayreh, W. Fan S. Gilbertson White, "Machine Learning Approaches to Predict Symptoms in People With Cancer: Systematic Review," *JMIR Cancer*, c. 10, e52322, Mar. 2024. doı: 10.2196/52322.

[64] M. L. Wallace ve diğ., "Use and misuse of random forest variable importance metrics in medicine: demonstrations through incident stroke prediction," *BMC Medical Research Methodology*, c. 23, no. 1, Haz. 2023. doı: 10.1186/s12874-023-01965-x.

[65] D. Ferreira-Santos, P. Amorim, T. Silva Martins, M. Monteiro-Soares P. Pereira Rodrigues, "Enabling Early Obstructive Sleep Apnea Diagnosis With Machine Learning: Systematic Review," *Journal of Medical Internet Research*, c. 24, no. 9, e39452, Eyl. 2022. doı: 10.2196/39452.

[66] A. W. Przybyszewski, A. Śledzianowski, A. Chudzik, S. Szlufik D. Koziorowski, "Machine Learning and Eye Movements Give Insights into Neurodegenerative Disease Mechanisms," *Sensors*, c. 23, no. 4, s. 2145, Şub. 2023. doı: 10.3390/s23042145.

[67] B. Sivajohan, M. Elgendi, C. Menon, C. Allaire, P. Yong M. A. Bedaiwy, "Clinical use of artificial intelligence in endometriosis: a scoping review," *npj Digital Medicine*, c. 5, no. 1, Ağus. 2022. doı: 10.1038/s41746-022-00638-1.

[68] J. Zhou ve diğ., "Predicting Stroke and Mortality in Mitral Regurgitation: A Machine Learning Approach," *Current Problems in Cardiology*, c. 48, no. 2, s. 101 464, Şub. 2023. doı: 10.1016/j.cpcardiol.2022.101464.

[69] N. Vakitbilir ve diğ., "Multivariate modeling and prediction of cerebral physiology in acute traumatic neural injury: A scoping review," *Computers in Biology and Medicine*, c. 178, s. 108 766, Ağus. 2024. doı: 10.1016/j.compbiomed.2024.108766.

[70] D. Amin, N. Garzon-Orjuela, A. Garcia Pereira, S. Parveen, H. Vornhagen A. Vellinga, "Artificial Intelligence to Improve Antibiotic Prescribing: A Systematic Review," *Antibiotics*, c. 12, no. 8, s. 1293, Ağus. 2023. doı: 10.3390/antibiotics12081293.

[71] H. Brand ve diğ., "High-Resolution and Noninvasive Fetal Exome Screening," *New England Journal of Medicine*, c. 389, no. 21, ss. 2014–2016, Kas. 2023. doı: 10.1056/nejmc2216144.

[72] F. Li ve diğ., "Application of artificial intelligence in VSD prenatal diagnosis from fetal heart ultrasound images," *BMC Pregnancy and Childbirth*, c. 24, no. 1, Kas. 2024. doı: 10.1186/s12884-024-06916-y.

[73] X. Liu ve diğ., "Applications of artificial intelligence-powered prenatal diagnosis for congenital heart disease," *Frontiers in Cardiovascular Medicine*, c. 11, Nis. 2024. doı: 10.3389/fcvm.2024.1345761.

[74] G. Tonni G. Grisolia, "Simulator, machine learning, and artificial intelligence: Time has come to assist prenatal ultrasound diagnosis," *Journal of Clinical Ultrasound*, c. 51, no. 7, ss. 1164–1165, Haz. 2023. doı: 10.1002/jcu.23512.

[75] L. Chitty ve diğ., "In case you missed it: The Prenatal Diagnosis editors bring you the most significant advances of 2019," *Prenatal Diagnosis*, c. 40, no. 3, ss. 287–293, Ocak 2020. doı: 10.1002/pd.5632.

[76] R. Arabi Belaghi, J. Beyene S. D. McDonald, "Prediction of preterm birth in nulliparous women using logistic regression and machine learning," *PLOS ONE*, c. 16, no. 6, P. B. Szecsi, ed., e0252025, Haz. 2021. doı: 10.1371/journal.pone.0252025.

[77] K. H. Nicolaides, "Screening for fetal aneuploidies at 11 to 13 weeks," *Prenatal Diagnosis*, c. 31, no. 1, ss. 7–15, Ocak 2011. doı: 10.1002/pd.2637.

[78] K. Spencer K. H. Nicolaides, "A first trimester trisomy 13/trisomy 18 risk algorithm combining fetal nuchal translucency thickness, maternal serum free $\beta$-hCG and PAPP-A," *Prenatal Diagnosis*, c. 22, no. 10, ss. 877–879, Eyl. 2002. doı: 10.1002/pd.420.

[79] K. H. Nicolaides, D. Wright, L. C. Poon, A. Syngelaki M. M. Gil, "First-trimester contingent screening for trisomy 21 by biomarkers and maternal blood cell-free DNA testing," *Ultrasound in Obstetrics and Gynecology*, c. 42, no. 1, ss. 41–50, Haz. 2013. doı: 10.1002/uog.12511.

[80] J. Yang, X. Ding W. Zhu, "Improving the calling of non-invasive prenatal testing on 13-/18-/21-trisomy by support vector machine discrimination," *PLOS ONE*, c. 13, no. 12, N. Shomron, ed., e0207840, Ara. 2018. doı: 10.1371/journal.pone.0207840.

[81] L. Botos, E. Szatmári G. R. Nagy, "Prenatal and postnatal genetic testing toward personalized care: The non-invasive perinatal testing," *Molecular and Cellular Probes*, c. 72, s. 101 942, Ara. 2023. doı: 10.1016/j.mcp.2023.101942.

[82] A. Jamshidnezhad, S. M. Hosseini, M. Mahmudi J. Mohammadi-Asl, "A machine learning technology to improve the risk of non-invasive prenatal tests," *Technology and Health Care*, c. 30, no. 4, ss. 951–965, Tem. 2022. doı: 10.3233/thc-213628.

[83] A. Koivu, T. Korpimäki, P. Kivelä, T. Pahikkala M. Sairanen, "Evaluation of machine learning algorithms for improved risk assessment for Down's syndrome," *Computers in Biology and Medicine*, c. 98, ss. 1–7, Tem. 2018. doı: 10.1016/j.compbiomed.2018.05.004.

[84] C. P. Weiner, M. L. Weiss, H. Zhou, A. Syngelaki, K. H. Nicolaides Y. Dong, "Detection of Embryonic Trisomy 21 in the First Trimester Using Maternal Plasma Cell-Free RNA," *Diagnostics*, c. 12, no. 6, s. 1410, Haz. 2022. doı: 10.3390/diagnostics12061410.

[85] M. Volk, A. Maver, L. Lovrečić, P. Juvan B. Peterlin, "Expression Signature as a Biomarker for Prenatal Diagnosis of Trisomy 21," *PLoS ONE*, c. 8, no. 9, Y. Herault, ed., e74184, Eyl. 2013. doı: 10.1371/journal.pone.0074184.

[86] S. I. Nagaoka, T. J. Hassold P. A. Hunt, "Human aneuploidy: mechanisms and new insights into an age-old problem," *Nature Reviews Genetics*, c. 13, no. 7, ss. 493–504, Haz. 2012. doı: 10.1038/nrg3245.

[87] T. Hassold P. Hunt, "To err (meiotically) is human: the genesis of human aneuploidy," *Nature Reviews Genetics*, c. 2, no. 4, ss. 280–291, Nis. 2001. doı: 10.1038/35066065.

# TEZDEN ÜRETİLMİŞ YAYINLAR

## Konferans Bildirisi

1. K. Türegün ve A. Yılmaz, "QF-PCR Sonuçları ile Prenatal Trizomi Riskinin Tahmin Edilmesinde Demografik, Klinik ve Genetik Verilerin Rolüne Makine Öğrenmesi Tabanlı Bir Yaklaşım," *13. Ulusal Moleküler Biyoloji ve Biyoteknoloji Kongresi (MOLBIOTECH-2024)*, Çevrimiçi, Türkiye, 1-3 Kasım 2024, (Sözlü Bildiri).